MAR 8 1944

Vol. XXXIV DECEMBER, 1943 No. 9

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

CONTENTS

Title-page and Index Volume XXXIV (1943) . . . . . i-viii

Is AQ or F Score the Last Word in Determining Individual Effort? . 513
FEI TSAO

Testing and the Course of Classroom Learning 526
ROBERT A. DAVIS

Eye Maturation and Reading Difficulties 535
GEORGE E. PARK AND CLARA BURRI

Eye Movements in Reading Type Sizes in Optimal Line Widths . 547
DONALD G. PATERSON AND MILES A. TINKER

Part and Whole Methods in the Memorizing of Music 552
CYRIL C. O'BRIEN

The Effect of Context on Learning a Vocabulary 561
C. L. MORGAN AND WILBERT L. BAILEY

Book Reviews 566

$6.00 per Year • Published Monthly September to May

WARWICK & YORK, INC.
BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md under the Act of March 3, 1879; additional entry as Second Class Matter at York, Pa.

THE JOURNAL OF Educational Psychology

Established 1910

EDITED BY

JACK W. DUNLAP
University of Rochester

IN ASSOCIATION WITH

STEPHEN M. COREY
University of Chicago

JOHN G. DARLEY
University of Minnesota

BERTHA PETERSON HARPER
University of Rochester

HAROLD E. JONES
University of California

H. H. REMMERS
Purdue University

PERCIVAL M. SYMONDS
Teachers College (Columbia University)

PAUL A. WITTY
Northwestern University

H. E. BUCHHOLZ
Managing Editor

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

Dr. Jack W. Dunlap is now on leave of absence from the University of Rochester and serving as Lieut. Commander in the U. S. Navy. For the 'duration', therefore, manuscripts, books and other materials for review, and correspondence regarding editorial matters should be addressed to H. E. Buchholz, Warwick & York, 10 E. Centre St., Baltimore, Md.

Manuscripts should be typed and double-spaced throughout including quotations, footnotes, and references. Return postage should be included with all unsolicited manuscripts.

THE JOURNAL is published monthly from September to May. The price per year in the United States and Pan-American countries is $6.00; $6.20 to Canada; and $6.40 to other foreign countries. Part-year subscriptions are 90 cents per issue ordered. Back volumes are $7.00 each, and back issues $1.10.

Subscribers should notify the Publishers of change of address at least four weeks in advance of publication of issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of next succeeding number.

WARWICK AND YORK *Publishers* BALTIMORE, MD.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXXIV December, 1943 Number 9

IS AQ OR F SCORE THE LAST WORD IN DETERMINING INDIVIDUAL EFFORT?*

FEI TSAO

University of Minnesota

Through a learning process a certain modification can be produced in the individual. The amount of gain of each pupil in learning depends upon his capacity to learn, his brightness or dullness. By a statistical procedure, taking the achievement score and either dividing by or subtracting the intelligence score, a measure of individual effort is obtained which is supposed to keep the function of capacity to learn constant.

Many methods have been used for estimating individual effort. All of them need two series of scores, one for intelligence and the other for achievement. In regard to the units, some measures of effort rating come from the age scales, others from T scales, and others are calculated directly from raw-point scales. Again, there are at least two classifications according to the computing procedure: one uses quotient or ratio technique, dividing the achievement score by the intelligence score; while the other uses the difference between the two functions. But all the measures are based on false assumptions regarding the correlation between intelligence and education. Moreover, each of them reveals a statistical fallacy. Therefore, they are not very useful estimates of individual effort. In the present paper, we mention only three measures; namely, accomplishment quotient(AQ), effort score(F), and efficiency ratio(ER), as the characteristic representatives of all the previously used methods.

* This study was carried out under the direction of the late Professor P. Sandiford and Dr. R. W. B. Jackson of the University of Toronto. Professor P. O. Johnson has given invaluable suggestions in dealing with the standard error of FQ. To all the author wishes to acknowledge his indebtedness. Most of the mathematical manipulations have to be omitted to conserve space.

The writer also suggests a new scheme in which both the actual correlation assumption and a clear-cut statistical theory are justified. The measure, which can be easily worked out step-by-step, is provisionally termed effort quotient(FQ). The relative merits and demerits of the suggested method and the ones used previously will be compared from both the practical and theoretical points of view.

THE DERIVATIONS OF VARIOUS MEASURES

The general process of computing an AQ consists in:

(1) Obtaining an intelligence quotient(IQ) or mental age(MA) on a mental test,

(2) Obtaining an educational quotient(EQ) or educational age(EA) on an achievement test.

(3) Using the formula for the AQ, i.e.,

$$AQ = \frac{EQ}{IQ}(100) = \frac{EA}{MA}(100)$$

The AQ technique has been widely used in educational research. The question of whether or not the AQ is sufficient to determine effectively each pupil's effort will be examined in the next section.

McCall's formula ([8], pp. 103, 276) for an F score is given below:

$$F = T_a - T_i + 50$$

where T_a is a T score on an achievement test; and T_i, a T score on an intelligence test. McCall's T unit is one-tenth of the standard deviation of an unselected sampling of twelve-year-old children([8], p. 278), but, for the purpose of experimental study, his unit may be in terms of the experimental group rather than twelve-year-old([8], p. 27). Moreover, the range of the T score is from 0 to 100 with an average of 50. F is not as popular as the AQ for estimating individual effort, but McCall and his followers use it enthusiastically. In his recent book, *Measurement*,[9] there is still sufficient room for a discussion of the measure of F.

Ford[3] proposed a tentative formula of effort rating which he termed efficiency ratio(ER):

$$ER = \frac{S}{Av\left(\frac{IQ}{100}\right)}(100)$$

where S is individual raw-point score made on an achievement test, and Av, the average of the class. Early in 1922, a similar idea had been presented by Torgerson[16], who called his measure the 'efficiency quotient.' But Torgerson, at the same time, pointed out that his method did not give as accurate results as might be secured with the use of age norms instead of grade norms. So he advised those who used the suggested method to keep in mind that the greater the deviation of the age of a given child from the average age of the grade, the more unreliable the measure becomes. Evidently the ER is greatly affected by the child's own age. This fact is easily seen when the ER formula is changed to the following:

$$ER = \frac{S(CA)}{Av(MA)}(100)$$

where CA denotes the chronological age and all the other notations are defined as before. Clearly, it is absurd to find that the older pupil tends to obtain higher ER points.

In view of imperfect correlation and unequal variability between an educational test and a mental test, the writer* suggests a new measure, which is provisionally termed effort quotient(FQ) and based upon the ordinary regression equation. The formula of FQ is:

$$FQ = \frac{E}{\text{Predicted } E}(100)$$

where E is an individual's educational score and predicted E is the predicted value of the educational score from his intelligence score.

The method of calculating FQ is as follows:

(1) Calculate the correlation between the intelligence scores and some achievement scores, denoted by r_{EI}. It does not matter whether the scores come from age scales, T scales, or even raw data.

(2) Calculate the standard deviations of these two, intelligence and achievement. These standard deviations are denoted by S_I and S_E, respectively.

(3) Denote by I, the known individual intelligence score; by M_I, M_E, the mean values of intelligence and achievement,

* Nygaard[13] and Cureton[2] had conceived the same idea. Their measures have been limited to the age scale, while the present scheme is extended to the T scale and even raw-point scale.

respectively; then the regression equation required to predict the individual educational score may be written:

$$\text{Predicted } E = M_E - b_{EI}M_I + b_{EI}I = b_{EI}I + K$$

where $b_{EI} = r_{EI}S_E/S_I$, and the value of K varies from one pair of intelligence and achievement tests to another.

(4) Substitute the value of E and predicted E in the above equation to find the FQ of the individual.

If a pupil's actual educational score equals his predicted educational score, the value of his FQ will be 100, which shows normal effort. Again, if his actual score is higher than his predicted score, the value of his FQ will exceed 100, which shows his diligence; and *vice versa*.

COMPARISON AND CRITICISM OF VARIOUS METHODS

All the scores of effort rating are denoted by quotient or difference measures computed from intelligence scores and achievement scores. We may examine, first, whether or not the relationship between scores on an intelligence test and an achievement test is perfect. From the previous studies([15,6,1], pp. 182–184, [5]) regarding this problem, the following conclusions may be reached:

(1) The correlations between the scores on the intelligence and achievement tests are positive, but far from perfect.

(2) The achievement tests differ in the degree of relationships, from very low to very high, between each of them and intelligence.

Now we may determine what is the rôle played by the correlation between education and intelligence in each of the effort measures. The measure AQ assumes that one whose EA or EQ score equals his MA or IQ score is normal as regards effort. Again, the normal F score corresponds to equal values of T_i and T_e. Similarly, in ER measures the standing of individual scores compared with the group norm must be adequately related to the IQ's. But the suggested FQ, on the other hand, is derived from the percentile ratio of the educational score to the predicted educational score, which is computed from the ordinary regression equation. It can be readily seen that the regression equation is based upon the actual correlation between the intelligence scores and the achievement scores, and the variability of the two sets of scores. In other words, the measures of FQ satisfy all the conditions, which differ from one pair of mental and educa-

tional tests to another, by using the regression equation. It follows, therefore, that the measures of AQ, F, and ER reveal at least the following defects:

(1) They assume that there are no differences from one subject to another in the degree of correlation. This assumption is not justified.

(2) They assume the correlation between intelligence and any subject to be +1.00. But this assumption is also not justified.

(3) All the measures but F compare directly the intelligence and the achievement scores, which may differ in relative variability (SD/mean). This introduces a statistical fallacy.

Furthermore, each of the following defects is special to one of these three measures.

(4) AQ cannot be calculated unless we have at hand the age scales for both intelligence and achievement.

(5) Theoretically, in a special case of $T_e < 25$ and $T_i > 75$, the value of F may be negative. It becomes meaningless.

(6) The value of ER is decidedly affected by chronological age. Since the older child tends to obtain higher ER points, it is clear this is not a satisfactory measure.

The writer carried out an experimental study to compare the merits and demerits of various measures. He asked a teacher in a public school in Toronto to give tests in Grade VII of his school. The materials used were the *Dominion Group Test of Intelligence* (Intermediate) of the Department of Educational Research of Toronto University and *Gates' Silent Reading Tests* Type D. The number of subjects available was sixty-five, divided into two rooms: one consisted of thirty-three, the other of thirty-two pupils. For each pupil we have raw score and age score, both of intelligence and reading. The results obtained are discussed below.

(1) *The correlations of RA with AQ, FQ and ER with the effect of MA partialled out.*—We calculated AQ and FQ from age scales, and ER. All of these measures are in terms of two rooms as a whole group. Using RA with the effect of MA partialled out as a criterion of effort, we find that:

$$r_{RA\ AQ \cdot MA} = 1.00$$
$$r_{RA\ FQ \cdot MA} = 1.00$$
$$r_{RA\ ER \cdot MA} = .84$$

It is also mathematically true that the correlation between RA and either AQ or FQ given by MA(or IQ) is approximately perfect. So, in any case, both the measures of AQ and FQ are adequate in estimating individual effort, at least as judged by the ideal criterion, RA given by MA(or IQ).

(2) *The correlation of IQ with AQ and FQ.*—Now we may examine the relationship between AQ and IQ. By mathematical manipulation we obtain the following approximation:

$$r_{AQ\,IQ} = \frac{r_{EQ\,IQ}V_{EQ} - V_{IQ}}{\sqrt{V_{EQ}^2 - 2r_{EQ\,IQ}V_{EQ}V_{IQ} + V_{IQ}^2}}$$

where V denotes SD/mean. Let us assume that $V_{EQ} = V_{IQ}$, and $r_{EQ\,IQ} < 1$. Then the correlation between AQ and IQ tends to be negative. No wonder that nearly all the previous studies show negative relationship between AQ and IQ, but this is by no means (as it might seem to be) the result of the psychological situation. Evidently, the negative tendency of $r_{AQ\,IQ}$ has already been fixed in the measures of AQ and IQ themselves.

The measure of FQ, however, is independent of IQ. By mathematical demonstration the quantity $r_{FQ\,IQ}$ is always approximately zero. In other words, the FQ is unaffected by the IQ. One who has a high IQ can obtain either a high or a low FQ. This is very fair from the psychological point of view. But with the AQ, the higher IQ is generally associated with the lower AQ. In criticizing Nygaard's revised AQ, which is synonymous with the FQ based on age scales, Odell([14], pp. 33–34) urges that the negative correlation between AQ and IQ is rather desirable in order to reveal the true psychological condition. But, unfortunately, the negative tendency of $r_{AQ\,IQ}$ comes not from any true conditions, but from the statistical computation. On this point, one may find little difficulty in comparing the relative merits of the FQ and the AQ, although both of them are perfectly correlated with an ideal criterion of effort rating.

(3) *The correlation between FQ and educational score.*—We find that the correlation between FQ and educational score is always approximately $\sqrt{1 - r_{EI}^2}$. The evident fact in favor of FQ is that the lower the correlation between educational and intelligence scores, the higher the relationship between FQ and educational scores; and *vice versa*. The scale of this approximation is given in Table I.

TABLE I.—RELATION BETWEEN r_{EI} AND $r_{FQ\ E}$

r_{EI}	$r_{FQ\ E}$
0.00	1.00
0.20	0.98
0.40	0.92
0.50	0.87
0.60	0.80
0.70	0.71
0.80	0.60
0.90	0.44
0.95	0.32
0.99	0.01
1.00	0.00

(4) *The constancy of various FQ's.*—Assuming that AQ can also be calculated from raw-point scales and T scales, the equations used in computing assumed "AQ_2" and "AQ_3" are as follows:

$$\text{"}AQ_2\text{"} = \frac{\text{raw-point reading score}}{\text{raw-point intelligence score}}(100)$$

$$\text{"}AQ_3\text{"} = \frac{T_e}{T_i}(100)$$

where T_e and T_i are both in terms of two rooms combined to form a single group.

Simultaneously, we calculated FQ_2 from two series of raw-point scores of reading and of intelligence, and their correlation in terms of two rooms combined; and FQ_3 from T_e, T_i and their correlation, again in terms of the whole seventh grade. Finally, we found the inter-correlations of various AQ's and of various FQ's as follows:

$r_{AQ\ \text{"}AQ_2\text{"}} = .70$ $\qquad$ $r_{FQ\ FQ_2} = .89$

$r_{AQ\ \text{"}AQ_3\text{"}} = .81$ $\qquad$ $r_{FQ\ FQ_3} = .90$

$r_{\text{"}AQ_2\text{"}\ \text{"}AQ_3\text{"}} = .93$ $\qquad$ $r_{FQ_2\ FQ_3} = 1.00$

Furthermore, we may examine the inter-differences among various AQ's and among various FQ's. Since "AQ_2" is too peculiar to be compared with the other two AQ's, it is not necessary to work out the inter-differences of "AQ_2" with the other AQ's. Thus, the results are reported in Table II, where each

figure given in the scale of the first column denotes the absolute difference, either positive or negative. From this table, nearly seventy per cent of the subjects show a difference between FQ and FQ_2 of less than 9 points; more than ninety per cent have a difference between FQ and FQ_3 of less than 9 points; and, again, nearly all the subjects have a difference of less than 9 points between FQ_2 and FQ_3. But on the difference of AQ and "AQ_3," sixty per cent of the children obtain a difference score of higher than 10 points.

TABLE II.—FREQUENCY OF INTER-DIFFERENCE AMONG VARIOUS FQ'S AND VARIOUS AQ'S

Difference Scale	AQ − "AQ_3" Frequency	FQ − FQ_2 Frequency	FQ − FQ_3 Frequency	FQ_2 − FQ_3 Frequency
0—4	18	22	32	46
5—9	8	22	28	18
10–14	10	13	2	1
15–19	11	5	2	0
20–24	9	2	1	0
25–29	5	0	0	0
30–34	4	1	0	0
Total	65	65	65	65

The very low inter-differences combined with the very high intercorrelations among various FQ's show the constancy of FQ. In other words, if the FQ can denote effectively the actual standing of individual effort, then it does not matter whether the FQ comes from age scales, T scales, or even raw-point scales.

(5) *Merits and demerits of the F score.*—We worked out F scores from T_s's and T_i's. Using RA, with MA held constant, as a criterion of effort rating as before, we found that

$$r_{RA\ F\ .\ MA} = .90$$

It seems that the measure of F is only a little better than ER, but worse than AQ and FQ as against an ideal criterion of effort estimate. But if we use T_s with T_i held constant as another ideal criterion, we obtain, in the present study,

$$r_{T_s\ F\ .\ T_i} = 1.00$$

It is not difficult to show, from algebraic manipulation, that this correlation must always be perfect. So the F score is also adequate to denote the estimate of individual effort. But we can also easily find that the correlation of F score with intelligence T score is always negative except when $r_{T_e T_i} = 1.00$. The relationship between $r_{T_e T_i}$ and $r_{F T_i}$ is shown in Table III, which

TABLE III.—RELATION BETWEEN $r_{T_e T_i}$ AND $r_{F T_i}$

$r_{T_e T_i}$	$r_{F T_i}$
1.00	0.00
0.90	−0.22
0.80	−0.32
0.70	−0.39
0.60	−0.45
0.50	−0.50
0.30	−0.59
0.10	−0.67
0.00	−0.71

comes from the approximation:

$$r_{F T_i} = -\sqrt{\frac{1 - r_{T_e T_i}}{2}},$$

which is always true. So the measure of F score is independent of the intelligence factor only if the correlation between achievement and mental tests is perfect. Since the relationship between achievement and intelligence is nearly always less than 1.00, the dull student generally tends to obtain a higher F score. No wonder that in McCall's recent book([9], p. 356), the expected result has been reported. But it may be explained in terms of only the F score itself. The smaller the correlation between intelligence and educational scores, the more does the measure of F favor the stupid children. This is naturally an important demerit of the F score.

Another defect underlying the F score has been mentioned before; that is, F score may be negative when $T_e < 25$ and $T_i > 75$. Of course, practically it may be rarely the case that one who has very high mental T score can obtain only very few T_e points. But if we consider the situation revealed by Table III carefully,

it can be readily seen that the case of a negative F value may happen when the correlation between Te and Ti is very low.

Besides its perfect correlation with an ideal criterion of effort, the measure of F possesses still other merits: (1) It makes the units of intelligence and achievement tests comparable. The standard deviation for either of these two tests is 10 points, while the mean is 50. (2) The mean value of F score in terms of the original group is always 50. This can justify the relative degrees of difficulty in achievement functions. But these merits, which can also be found in the measure of FQ, cannot entirely compensate for the demerits of F score.

(6) *The distributions of various measures in the present group.*— We found that the distribution of each FQ series in terms of the whole experimental group is concentrated around 100 points as a mean value. Nearly half of the subjects have each FQ of more than 100; while another half, less than 100. Except for the F scores and the assumed "AQ_3" scores, both of which come from T scales, all the other series of effort ratings have an average of less than expected value. In this case, it is also theoretically true that the mean FQ of any group will be 100, and the average F will be 50, irrespective of various difficulty degrees of different achievements. But the average value of AQ or ER is uncertain. In the present study, it seems that most of the subjects tend to be lazy according to the measure of AQ or ER. But in other cases, when the achievement test is easier, most of the pupils, including lazy ones, will obtain high AQ or ER points. So these two measures are by no means useful according to this consideration.

Odell([14], p. 33), however, pointed out that if any group has an average of 100, then the situation of very poor instruction in some classes cannot be revealed. This is, of course, a very important point to be taken into consideration. But it can be easily settled. When we wish only to compare the standing on effort of each pupil within a certain group, we compute the FQ simply from this group. When, on the other hand, we wish to rate the average effort of several classes, we may work out the FQ in terms of these classes as a whole group. Thus, we can easily find that the average FQ of Room A is slightly lower than that of Room B in the present study. As other factors are not equal, however, we cannot conclude from this result that one teacher is inferior to the other.

(7) *A discussion of the standard error of the FQ**.—We know that the mean standard error of the predicted educational scores from intelligence scores is shown in the following form:

$$\sigma_{E.I} = S_E \sqrt{1 - r_{EI}^2}$$

where $\sigma_{E.I}$ denotes the average standard error of a predicted educational score from an intelligence score and all the other notations are defined as before. Let us assume that

$$t = \frac{\text{mean value of predicted E}}{\sigma_{E.I}}$$

or

$$\sigma_{E.I} = \frac{\text{mean value of predicted E}}{t}$$

Then, from mathematical manipulation, the mean standard error of the FQ is approximately FQ/t. The smaller the $\sigma_{E.I}$, the closer the approximation. Furthermore, if we deal with the standard error of each individual FQ score, the same result can be obtained. In other words, the ratio of an individual FQ score to its own standard error is also approximately equal to the ratio of the corresponding predicted E score to its own standard error. Once again, the smaller the standard error of the corresponding predicted E score, the closer the approximation. It follows that the significance of the FQ scores is nearly the same as the significance of the predicted E scores. From the above formulae the following statements may be made: (1) The closer the relationship between educational and intelligence scores, the more reliable the FQ. (2) The smaller the variability of educational scores, the more reliable the FQ.

It is hard to say what is the sufficient degree of relationship between educational and intelligence scores. If the variability of educational scores is very small, we would have a fairly reliable set of FQ scores whether or not the correlation between educational and intelligence scores is high. If, on the other hand, the variability of educational scores is very large, then the relationship between achievement and intelligence scores should be very high in order to obtain a reliable set of FQ's. But in any case one will have no difficulty to evaluate approximately the standard error of FQ scores to show how reliable they are.

---

* This discussion is sufficient for the practical purpose. A more rigorous investigation of this point will be carried out later.

SUMMARY

(1) FQ, AQ, and F are all adequate in determining the pupils' relative positions on effort as against an ideal criterion.

(2) ER cannot be regarded as a perfect measure of effort as against an ideal criterion.

(3) AQ, F, and ER are based on false assumptions regarding the correlations between intelligence and achievements, while FQ considers the true facts.

(4) With any two series of scores, one in intelligence and the other in education, FQ, ER, and F can be easily calculated; but AQ cannot be worked out unless we have at hand the age scales both of intelligence and achievement.

(5) It is possible that a pupil may have a T_1 of more than 75 and a T_2 of less than 25. In such an extreme case, the value of his F will be negative.

(6) The measure of AQ or ER compares directly two series of scores, whose relative variability (SD/mean) are unequal; while the measure of FQ or F puts any two series into equal units.

(7) The mean value of FQ within a group is always 100, and of F, 50; but the average of AQ or ER is not fixed.

(8) Taking various FQ's in same group to be compared, their intercorrelations are very high, and their inter-differences are very low. This demonstrates the constancy of FQ.

(9) The correlation of AQ or F with the intelligence function usually tends to be negative, whereas FQ is always independent of the intelligence factor.

(10) The older pupils tend to obtain higher ER points.

(11) The relationship between the FQ and education is dependent on the degree of the correlation between intelligence and education. The lower the correlation between intelligence and education, the closer the relationship between the FQ and education, and *vice versa*

(12) The relative size of the standard error of an individual FQ score is approximately the same as the relative size of the standard error of the corresponding predicted educational score.

CONCLUSIONS

(1) ER is decidedly not a good measure of effort.

(2) AQ and F may be used in determining only the pupils' relative positions on effort.

(3) FQ can be used as a measure not only of the pupils' relative positions but also of their absolute standings on effort.

BIBLIOGRAPHY

1. Burt, C. *Mental and Scholastic Tests.* (4th edition) London: P. S. King, 1933, pp. xv + 632.

2. Cureton, E. E. "The Accomplishment Quotient Technic," *Journal of Experimental Education*, v, 1936, pp. 315–326.

3. Ford, F. A. *The Ratio of Achievement to Ability as Found among Fifth-grade Pupils.* Contribution to Education, No. 94, George Peabody College for Teachers, Nashville, Tennessee, 1931, pp. 61.

4. Franzen, R. H. "The Accomplishment Quotient of School Marks in Terms of Individual Capacity," *Teachers College Record*, XXI, 1920, pp. 432–440.

5. Gates, A. I. "The Correlations of Achievements in School Subjects with Intelligence Tests and Other Variables," *Journal of Educational Psychology*, XIII, 1922, pp. 129–139, 223–235, 277–285.

6. Hollingworth, L. S. *The Psychology of Special Disability in Spelling.* New York: Teachers College Contribution to Education, No. 88, 1918, pp. vi + 105.

7. Jackson, R. W. B. "Some Pitfalls in the Statistical Analysis of Data Expressed in the Form of IQ Scores," *Journal of Educational Psychology*, XXXI, 1940, pp. 677–685.

8. McCall, W. A. *How to Experiment in Education.* New York: Macmillan, 1930, pp. xiv + 281.

9. McCall, W. A. *Measurement.* New York: Macmillan, 1939, pp. xv + 535.

10. Monroe, W. S. *The Illinois Examination.* University of Illinois Bulletin, XIX, No. 9, 1921, pp. 70.

11. Monroe, W. S., and Buckingham, B. R. *Illinois Examination: Teacher's Handbook.* Urbana: University of Illinois, Bureau of Educational Research, 1920, pp. 31.

12. Monroe, W. S., and Buckingham, B. R. *The Illinois Examination I and II: Teacher's Handbook.* Bloomington, Illinois: Public School Publishing Co., 1920, pp. 32.

13. Nygaard, P. H. "A Revised Accomplishment Quotient," *Journal of Educational Research*, XVIII, 1928, p. 87.

14. Odell, C. W. *A Critical Study of Measures of Achievement Relative to Capacity.* University of Illinois Bulletin, XXVI, No. 29, 1929, pp. 58.

15. Proctor, W. M. "The Use of Intelligence Tests in Educational Guidance," *School and Society*, VIII, 1918, pp. 473–478, 502–509.

16. Torgerson, T. L. "The Efficiency Quotient as a Measure of Achievement," *Journal of Educational Research*, VI, 1922, pp. 25–32.

TESTING AND THE COURSE OF CLASSROOM LEARNING[1]

ROBERT A. DAVIS

University of Colorado

Broadly speaking, every evidence of learning is a test of some kind; and every result of testing is evidence of learning. The learner during his study is subjecting himself continuously to testing procedures, whether he studies a subject that lends itself to fixed answers or one that involves controversial issues in which answers are a matter of opinion. The learner is at all times employing self-testing exercises either through a process of recall or performance in response to learning materials. Self-testing in some subjects such as mathematics, spelling, and foreign languages is obvious; but any subject lends itself to evaluation. The learning process in the more formal classroom situation, however, consists in providing for a period of instruction and study, followed by testing, which in turn is followed by further instruction and study, and so on. Whether the intimate relationship of these two processes is provided for in teaching depends upon one's formulated philosophy of teaching and upon one's willingness to prepare for systematic learning and testing before instruction begins.

Minimum essentials for systematic study of the course of learning for a given school subject demand careful preparation by the teacher of two closely related steps in planning: (1) formulating significant objectives and devising test situations which will measure the extent of progress toward their attainment; and (2) the administration and uses of tests with a view to the advancement of learning. The first step requires a systematic plan for learning and testing; the second requires putting the plan into operation. There is, of course, an intervening step which involves the cultivation of the abilities sought as teaching aims. This important step, however, is outside the province of this paper.

---

[1] Some of the principles presented in this paper have already been published for local use by R. T. Purnell and R. A. Davis, *Directing Learning by Teacher-Made Tests*, Bureau of Publications, Extension Division, University of Colorado, 1939.

Responsibility for formulating teaching objectives rests upon the teacher who, in planning and developing his course, must decide which mental processes or types of behavior he wishes to cultivate through the teaching of his subject. He is confronted with the very important problem of deciding how thoroughly he wishes his pupils to learn, how deep and broad their knowledge should be, and the type of attitudes or beliefs it is desirable to develop. He must decide whether he wishes his pupils to be able to recall and recognize only the essential factual material of his subject; or whether, in addition, he wishes them to be able to apply principles[1] learned, to interpret factual material, to formulate logical conclusions from given facts, or to acquire certain attitudes or beliefs. Some of these abilities may be cultivated in many types of subject-matter, whereas others may be applicable only to certain fields. In any case, he will find it desirable to select a few significant objectives most amenable to measurement, and strive for their development. These objectives, which should be acceptable to pupils (if not actually formulated in coöperation with them) and made available in mimeographed or other form for guidance, serve to make instruction and study purposeful during the course of learning and definitely assist pupils in preparing for tests and examinations.

Testing situations, to be effective in directing the course of learning, must be specifically related to teaching objectives. Teachers talk glibly about types of performance required in their subjects, but in practice stress the ability to reproduce or recognize factual information. This ability represents the most typical result of learning (memory). In its extreme form, it is rote memorization, which is learning without comprehension. In whatever form it is found, however, the ability to draw on memory has been the objective most frequently stressed in teaching and in testing.

Two situations account for this extraordinary emphasis upon the ability to reproduce or recognize factual information. One is the notion, concerning which there is little disagreement, that factual information is the raw material on the basis of which

[1] The Evaluation Committee of the Eight-Year Study has demonstrated the desirability of devising tests to measure the learner's reactions in many situations. See Smith and Tyler, *Appraising and Recording Student Progress*, Harper's, 1942.

other abilities must be developed. The fallacy in this notion is in considering factual mastery as an end rather than as a means to the development of other abilities. The other is that, although teachers may stress the development of other abilities, they have followed in their testing the path of least resistance and have used either standardized tests or their own quickly and easily constructed tests. It is perhaps not an overstatement to say that the typical standardized test and conventional teacher-made test measure primarily recall and recognition of factual information.

After objectives have been formulated and test situations devised, the teaching program may be launched. At this point, two closely related problems of planning present themselves: (1) administration of tests at appropriate periods during the course of learning; and (2) the uses or purposes to be served by tests. The place of tests in the course of learning as well as the frequency of their administration obviously depends upon the uses to be made of the results. The effectiveness of any program depends upon the extent to which test results reflect objectives and the patient, thorough teaching to reach those objectives.

Programs for testing make use of plans which may be classified as follows: (1) tests which are given before formal instruction and study begin; (2) tests which are given at any time on material which has been studied between the beginning of a subject and its completion; and (3) tests which are administered periodically over broad fields of subject-matter after the completion of individual subjects. The first two plans are definitely related to the course of learning, whereas the third assumes a testing program which is as continuous as learning itself.

A desirable use of a test is to survey at the beginning of a subject the pupil's previous background and the extent to which any abilities have already been developed. Certain work skills and habits may also be appraised through this initial testing. The results of the initial testing may be compared with those of a testing period at some later date, as, for example, at the completion of a subject, in order to determine the extent of progress in development of the abilities tested. Initial testing results may be used also as a guide to selection of teaching methods and to special provisions for individual needs. The test may serve also to clarify misconceptions of pupils regarding reasons for use of tests and to create a coöperative attitude between teacher and pupil.

Similarly, the plan may be used at the beginning of either a specific assignment or of a major unit of material as a means of determining needs of pupils. One form of a test may be used for initial testing and a comparable form to measure the results of learning. The pre-test can be used profitably wherever subject-matter is such as to suppose an indeterminate amount of previous knowledge or incidental learning. It may inform the pupil wherein he is weak or strong, thus suggesting the need of more study in some aspects of subject-matter and less in others. If the pupil knows that a pre-test is given for guidance and not for measuring accomplishment, its psychological effect is good.

If testing be regarded as an integral part of learning, it follows that tests may be given at any time from the beginning of a subject until its completion. Tests so given should indicate the extent to which teaching objectives are being attained, as well as the nature and degree of error in learning. Depending upon the time of their administration and the amount of subject-matter covered, they may be legitimately used during the course of learning both as a basis of guidance and as a measure of accomplishment. There is, therefore, no rule which would apply validly to the place for testing in the course of learning or the frequency with which tests should be given. Practical considerations, however, demand that we be more explicit with regard to the time and place for administering tests. Four plans for testing during the course of learning, therefore, will be briefly described: (a) testing to accompany specific assignments; (b) testing after the completion of a comparatively large unit of subject-matter; (c) testing at an arbitrary period such as midterm; and (d) testing over the entire range of material at the end of a subject, commonly known as the final comprehensive examination.

TESTING TO ACCOMPANY SPECIFIC ASSIGNMENTS

Testing may be planned to accompany specific assignments. This is usual in elementary- and junior-high-school grades and less common as the higher grade levels are reached. Even though the teacher may have organized his work into major units, there may be need for checking frequently progress in mastering specific assignments. Tests used to accompany specific assignments are thus intimately related to learning; they are used primarily for the purpose of analyzing pupil difficulties, of measuring growth, and for providing the teacher with materials

for guidance. It is recognized, of course, that in some subjects such as mathematics and foreign languages each class meeting may be a testing period in that pupils are required to answer questions, solve problems, and demonstrate their knowledge in some objective manner. But in other subjects in which wide reading may be required testing may take the form of short written tests.

The plan has the advantage of accustoming pupils to types of test situations which may be used later, of removing misconceptions, of directing attention to the most significant material, and of making it possible for the learner to secure immediate recognition of his effort in completing specific assignments. It also provides the teacher with an abundance of pupil reactions to situations related to teaching objectives. These reactions may be re-formulated and subsequently used in preparing more comprehensive tests and examinations. Teachers may also require that pupils formulate answers to short tests, as homework, frequently assigning textbook problems or special questions demanding types of performance sought as the aims of the course.

In general, the short tests given during the class period and homework exercises should be regarded as teaching material. If this be done, pupils will feel assured that the purpose of the work is training and that, although they will be informed of their proficiency, their ultimate record will not be greatly affected, provided that their performance be consistent. The burden of correcting short tests and written exercises may be shared by pupils in scoring the papers. Obviously, however, pupils cannot score papers unless items are of the objective type. Tests following specific assignments probably come closest to fitting into the ideal teaching program because they provide immediate learning situations for the learner. The situations should properly represent all teaching objectives formulated for the subject, a sufficient quantity of items being provided at one time or another for measurement of each objective.

TESTING AFTER THE COMPLETION OF EACH MAJOR UNIT OF SUBJECT-MATTER

Wherever subject-matter is organized into broad meaningful units embracing closely related and interdependent materials, whether by textbook or by specially prepared syllabus, we have

a logical basis for administering tests. Tests given on material of such a unit constitute a sound basis for determining difficulties and needs of pupils before beginning a new unit of subject-matter upon which subsequent units may depend. These tests, like those accompanying specific assignments, should be used primarily for instruction and guidance, even though they may be sufficiently comprehensive to be used as a measure of accomplishment. They serve to identify learning difficulties, to reveal the nature and extent of errors in learning; and they afford a basis for remedial instruction inasmuch as difficulties are revealed to both teacher and learner, as the course of learning proceeds from unit to unit. Such tests have the effect of causing pupils to think of materials as a whole, and not as disconnected bits of subject-matter. They also harmonize with systematic reviews which are usually scheduled at the completion of a major unit of material.

Although the number and size of units may be expected to vary with types of subject-matter, in general the larger the unit the greater the economy of learning and of testing. Because of the comprehensiveness of material of most units, however, it is possible to construct tests of such length as to possess satisfactory reliability and validity. The test results may, therefore, be used as a basis of appraisal, in addition to their instructional value. Each major unit of work, inasmuch as it presupposes some statement of objectives and psychological organization, represents one of the most promising bases for testing during the course of learning.

TESTS GIVEN AT THE HALF-WAY MARK—'MID-TERM TESTS'

In many schools, especially the colleges and universities, the practice has grown up of giving a comprehensive test over all of the work covered up to the half-way mark of a quarter or semester. Principals of high schools and deans of colleges and universities like to check the student's progress at the half-way mark so that he may be warned in due time of possible failure. High-school officials also use the results of the mid-term test as a basis of preparing report cards for parents. The practice appears to be based on the questionable assumption that comprehensive test results are needed to indicate the extent to which the pupil is making satisfactory progress and to reveal weakness in his

learning before it is too late to improve his status. In some cases the practice may be justified, but the teacher should certainly know whether a pupil is making satisfactory progress before the middle of the term. If tests be administered over specific assignments and over major units of material, the chances are that the teacher has sufficient knowledge of the needs of a learner for instruction and an adequate measure of his accomplishment. This is not to say that comprehensive tests may not be given at some stated period; but the practice of giving them merely at some arbitrary period in the course of learning cannot be justified as rationally sound for instruction or psychologically sound for learning.

THE FINAL COMPREHENSIVE EXAMINATION

In the light of the preceding discussion, it is apparent that the final comprehensive examination should represent one more link in a chain of testing during the course of learning. It may be justified only insofar as it affects the advancement of learning and serves as an adequate basis of appraisal. The short tests accompanying specific assignments and the longer ones covering major units of material serve as stages in learning and reach their culmination in the final comprehensive examination, which should require the learner to demonstrate more fully and perfectly than has hitherto been possible the knowledge that he has gained during all of his preceding instruction and study.

The final comprehensive examination provides opportunity for intensive and extensive sampling of learning material, and thus furnishes a sound basis of rating. It measures the extent to which the pupil has retained his knowledge, particularly that gained during the early and middle stages of learning the subject; consequently, it furnishes evidence concerning the thoroughness of learning. Opportunity is provided for the teacher to measure the degree of attainment in the objectives formulated at the beginning; the short quizzes and tests preceding merely indicate stages in their development. The final examination (if it be understood that it will be required) stimulates the learner to study his material as a whole, to synthesize his knowledge, to form a chain in his thinking—to make him keep in repair the entire course of learning as he goes along.

Investigations dealing with the effect of exemption from the final comprehensive examination, the extent to which it provides information additional to that the teacher already has by the time it is given, have not yielded answers sufficiently conclusive for generalization. Some studies,[1] however, indicate that in situations in which testing at regular intervals during the course of learning has been arranged the total cumulative initial acquisition scores are very closely related to the final examination scores. Furthermore, if pupils be remeasured at the end of a two-week period by means of duplicate forms of initial acquisition tests during the course of learning, their total cumulative retention scores (two-week period) correlate even more closely with final examination scores based upon duplicate forms of the same tests. It may be concluded, therefore, that the best assurance for superior performance on the final comprehensive examination is a substantial amount of initial acquisition and its retention (at least for a short period) during the course of learning. Certainly, those are minimum requirements for thorough learning and good retention. 'Cramming' may have the effect of producing immediate results if the examination requires reproduction of a body of factual material, but it has little virtue as a study-practice in preparation for tests and examinations which demand higher levels of thinking.

The final comprehensive examination may be similar in structure to previous tests except that it need not stress analytic and diagnostic provisions. It may be expected to reveal, if its results are compared with those of previous tests in the subject, progress toward attainment of selected objectives. Although its results cannot be utilized as a basis for further instruction, they nevertheless make it possible for learner and teacher to judge the effectiveness of learning and instruction in attaining their objectives.

Testing is as continuous and necessary as learning itself. Perhaps the principal criticism that may be made of learning and of testing in school and college is the stress upon numerous courses restricted to a quarter or a semester and the recognition of their completion by the administration of a final comprehensive

[1] A. H. Word and Others, "Acquisition and Retention of Factual Information in Seventh-grade General Science during a Semester of Eighteen Weeks," *Journal of Educational Psychology*, 1939, xxx, 116–125.

examination. The practice, at best, encourages the pupil to learn superficially only a small sample of a broad field and to forget what he has learned as soon as the course has been completed. But even though the present system be retained, testing may be planned to serve a wider and more useful purpose.

Testing programs should be planned to include broad fields of subject-matter irrespective of specific courses in the learner's schedule. If comprehensive tests are designed for the measurement of higher abilities as well as the ability to reproduce and recognize factual information, they will indicate the extent to which the pupil is growing in depth of knowledge and understanding. They should be given principally for guidance and thus made to serve as a constant stimulus for growth. If tests be given for the purpose of stimulating improvement, the learner will enthusiastically indorse their use. Studies dealing with retention in school subjects indicate that although specific factual information is illusory and quickly forgotten, higher abilities such as the ability to apply principles, to generalize, to interpret data, and to solve problems are retained over long periods with slight loss.[1] Testing over broad fields would indicate the effectiveness of certain teaching objectives stressed in particular courses and would measure the extent to which the learner is gaining new knowledge and deeper insight as a result of his school experience and increased maturity. Most of all, continued testing would indicate the nature and extent of the learner's permanent knowledge and the proficiency he has gained in applying it.

[1] A. H. Word and Others, "Individual Differences in Retention of General Science Subject-matter in the Case of Three Measurable Teaching Objectives", *Journal of Experimental Education*, 1938, VII, 24–31.

R. W. Tyler, "Permanence of Learning", in "With the Technicians", *Journal of Higher Education*, 1933, IV, 203–4.

EYE MATURATION AND READING DIFFICULTIES

GEORGE E. PARK AND CLARA BURRI

Department of Ophthalmology, Northwestern University and Dyslexia Memorial Institute, Wesley Memorial Hospital

It is a commonly accepted fact that the process of reading, at first sight, apparently is so simple that all which is required is to have stimuli presented and a reaction to the stimuli. However, it is the complexity of these two aspects, i.e., of getting and responding, which reveals the great problem in the interpretation of any visualized object, be it a whole situation in which many objects appear or a distinctive field, a part of a field, or a sign or a symbol. Hence to interpret reading and reading failure one must constantly be aware of this complexity.

The optimum in visual perception for efficient reading consists not only of good binocular visual acuity, but also a relative normal status of all the various factors concerned in the whole visual process. It is possible to have good reading habits with monocular fixation, but more often reading difficulty may develop if both eyes are functioning with poor coördination so that fusion is accomplished with difficulty. This has been shown to occur at any stage of development in two previous papers.[1,2] In this report particular attention is directed to the observations that reading difficulty may have its origin in a delayed maturation of the sensory-motor responses to visual stimuli. So when a child undertakes to learn to read before his sensory impulses or his motor responses are fully matured, he may become a poor reader.

One problem was to determine at what age a child's eyes mature sufficiently for him to learn to read efficiently. In other words, when is a child ready to read? Can it be assumed that every child will have fully matured visual development when he enters school at the age of six? Isn't it conceivable that some children are not yet ready to take up successfully the complex task of learning to read because their eyes are not fully developed?

At birth and shortly after, there is little coördination of the eyes and, therefore, limited simultaneous binocular vision. The child must learn how to use his eyes together, and the process of learning binocular vision differs little from any other learning process.[3] At first a child exercises his muscles by trial and error

and finally by control of these random movements clear vision is achieved. Gradually he learns to repeat the movements which, at first more or less accidentally bring clear and single vision. Eventually he learns to associate a certain muscle pattern with clear single vision and then integrates his eyes accordingly. Through the repeated reciprocal innervation and inhibition of the opposing muscles the two images tend to fall on the foveae, and through constant repetition the child learns to fuse the separate images.

For normal vision it is necessary to have vision acuity of 20/20 or better either uncorrected or corrected by refraction, good fusion with adequate duction reserve and good stereopsis. Binocular vision comprises three principal processes; namely, 1) the coördination of the extraocular muscles and the intrinsic adjustment of the ciliary and other internal muscle fibers in order to bring the two images on the foveae; 2) the fusion of the two monocular images or sensations into a single perception; 3) the individual's conception of the visual stimuli.

Fusion is an association of the images received by the two retinas so that one composite picture is formed. The images must be fused not only for single vision but also for a 'true impression of depth,' or stereopsis. In order that fusion may take place effectively the extraocular and intrinsic muscles must maintain the proper tension and tonicity to cause the two images to be focused on the foveae.[4,5] The process is gradually learned throughout the first few years of life. It may be altered or perfected throughout the greater part of life, or changes may be made as long as one is capable of learning.

But before a child has fusion he builds up the 'accommodation-convergence' association. This occurs through association of movements of convergence and the innervation of the ciliary muscles, which are purely proprioceptive processes. At this stage the child's visual behavior is mainly a proprioceptive space direction affair, with poor coördination which may result either in over-convergence or in over-innervation of his ciliary muscles. This causes diplopia and blurred vision, which the child gradually recognizes from his association with touch. He discovers that when his eyes tell him there are two objects, his hands tell him there is only one. This stage is probably reached at the age when clearer and more accurate vision is needed. Then a higher

integration is developed. Through voluntary attention he learns to control indirectly his internal recti muscles and his ciliary muscles, because of their associated function with convergence. Thus, by means of voluntary direction of attention he adjusts his eyes until clear vision is obtained and he has fusion. The fusion becomes his interpretation or perception of a specific image and the response pattern of his visual impulses. Thus, through the first years of life a child gradually learns to use his two eyes together.

At present there is available little factual material which throws light on the development of binocular vision in early childhood. There are only isolated studies on blinking, pursuit, and eye movements during infancy. In the first year of life, a baby's eyes show little coördination and accordingly little simultaneous binocular vision. Not much is known of how and when full binocular vision and fusion and stereopsis are developed.

Contraction or dilatation of the pupil seems to be the only ocular response which is fully present at birth. During the first day it may be slow and inconstant, but by the second day it is fairly well established.

The movements of lids and eyeballs of the infant are frequent but very incoördinated during the early period of life. Several observers report the closing of one eye while the other eye is open; also the partial closing of eyes when awake and asleep. Although the blinking reaction usually occurs within a few hours after birth, the quick blinking which occurs in adults at the approach of an object is not present in infants before fifty or sixty days, according to Preyer[6] and Blanton.[7]

Reports concerning coördinated ability to follow moving objects vary considerably. Shinn[8] thought that until nearly three months her niece had not really "looked at any special objects which were not lighted up, colored, or in motion." Jones[9] in a study based on more than three hundred subjects showed that horizontal eye-following movements appeared before vertical, and vertical before circular. Half of her subjects followed a light horizontally by the thirty-eighth day, all by the ninetieth day. Ocular pursuit of rapid movement comes appreciable late and by slow stages. A searching regard for objects after disappearance also matures gradually, but most children look after a fallen spoon at forty weeks.

Ling[10] studied visual fixation in twenty-five infants varying in age from seven minutes to twenty-four weeks, utilizing a circle in the center of a white cylinder. A two-inch black stimulus disk was made slowly to recede from and approach the child's eyes within a distance of three feet. Motion pictures and written records of each child's behavior showed that sustained visual fixation, while absent at birth, first appears within a few hours, reaching a peak at four to five weeks. Binocular fixation accompanied by convergence does not appear until seven to eight weeks of age.

Aside from these few studies on blinking, visual fixation and eye movement during the first year of life, no controlled observations on the development of binocular vision exist. There is particularly an absence of controlled studies of the genetic development during the second and third years of life. We know little about the development of visual acuity, fusion and stereopsis in children below the age of six. Johnson and Bech[11] attempted to study stereopsis in a more or less naturalistic setting in which the child could reveal the presence of stereopsis through his spontaneous reactions to the situation. Twenty subjects, two to six years old, viewed through polarized glasses two superimposed, polarized colored pictures of a doll projected to life size on a ground glass screen. To an adult with normal binocular vision this set-up produced the stereoscopic illusion that the doll was between the observer and the screen. All children, irrespective of age, reached for the doll several inches in front of the screen, and from this and the verbal reports the authors concluded that stereoptic vision is present at two years of age.

However, other observers have been impressed by the absence of stereopsis in many children as old as five and six years. Mann[12] believes that stereopsis is present at the age of six. Of course, great individual differences may be expected, some children having fusion and stereopsis much earlier.

Lancester[13] reported that the visual acuity of children on the whole, is lower than that of adults, and Nugent and Ilg[14] found that many five- and six-year-old children whom they tested failed to have fusion and stereopsis, especially at the near point. That observation agrees with the findings of this study. It was found that a great many of the first-graders had vision below 20/20 even though they had little refractive error, and that

many had poor fusion and little stereopsis. The problem is apparently one of maturational retardation or low achievement in the development of good binocular vision, and not one attributable to visual defects.

In order to discover whether this poor binocular function in the pre-reading and beginning reading groups was consistently greater than among the older children, an analysis of the frequency of the various eye factors was made and compared with those in the older age groups. It was found that there were consistently more students with poor vision with little or no fusion or stereopsis, and poor duction ability among the younger children than in the other age groups.

In order to make this comparison, the data from the two hundred twenty-five school children who were tested for various eye defects and reading difficulties were analyzed in terms of the effects of these various eye defects on reading difficulties. The procedure of these examinations and their results have already been reported in two previous papers.[1,2] Therefore, a full description of the method of examination and analysis will not be repeated. It will suffice to state that the subjects of this study were a group of two hundred twenty-five unselected Chicago city public school children, ranging from the first to the eighth grade, and that each child received a relatively complete eye examination, consisting of a vision test with and without cycloplegic, a homatropine refraction, and also phoria, duction and fusion tests for twenty feet and for the reading distance.

In order to obtain a comparative score for eye defects, each eye condition was arbitrarily assigned a value of one point. Thus, the more defects the higher the score. Zero represented a completely negative examination, while 14 was the highest and poorest score possible.

After completing the eye examinations the school furnished us the records from the City of Chicago Reading Tests and the mental age scores for every child. Thus, a given child's reading accomplishment and intelligence quotient were not known to us at the time of the eye examination.

The reading scores reported in our study are in terms of mental age expectancy. Such scores were obtained by dividing the reading score,* in terms of grade expectancy, by the grade level

* From the Chicago Reading Tests.

in terms of the mental age. Thus we attempted to measure reading achievement in terms of what an individual could be expected to do according to his ability or intelligence and not just according to his grade level.

In general it was found that abnormal peripheral ocular variations are concomitant in a direct ratio with low reading skill and inefficiency, and that certain of the eye factors have a greater relationship to reading difficulties than others. Binocular disturbances are the ones most frequently related to low reading achievement; even more than refractive errors. As a rule monocular individuals have little trouble with reading insofar as their eyes are concerned. Absence of simultaneous binocular vision gives a small and insignificant correlation between reading scores and eye abnormalities. It is chiefly interference resulting from abnormally functioning binocular relationship that causes interference in the process of reading. If for some reason or other, the neuro muscular mechanism of the eyes is disturbed so that they can no longer function efficiently as a unit, fusion has a tendency to break, causing potential diplopia, and the resulting interference becomes a disturbing factor in the process of reading.

If this is possible, and affects reading ability in later periods of life, how much more important such interference from lack of binocular vision will be if it occurs during the time when the child is just beginning to learn to read. For instance, if a child is forced to do much close work, as often happens with children who are highly intelligent or who have overly anxious parents, before the binocular mechanism is properly matured, bad binocular habits may be established, which in turn may react detrimentally upon the child's desire to learn to read. Children who are urged to do much close work before normal binocular vision is developed may also develop psychological manifestations. These quickly enter the picture when a child fails in normal scholastic achievement. Emotional 'sets' against teachers and parents and the school in general are frequent complications.

In view of these possibilities it is of prime importance that a child's visual apparatus be mature at the time he starts to learn to read. It has been customary to consider a child ready to learn to read when he is six years old. But today many educators

TABLE I.—EYE MEASUREMENTS OF PRE-READERS IN RELATION TO READING READINESS (RR) AND GRADE EXPECTANCY IN TERMS OF MENTAL AGE

Subject	Vision S.		Vision C.		Phoria (Distance)		Phoria (Near)		Duction (Distance) Abduction		Duction (Distance) Adduction		Duction (Near) Abduction		Duction (Near) Adduction		Fusion	Stereopsis	Reading Readiness	Mental Age Grade Expectancy
	R	L	R	L	Exo-	Eso-	Exo-	Eso-	Break	Return	Break	Return	Break	Return	Break	Return				
1	20/20	20/20	.....	.....	0	0	0	0	6	2	6	2	9	6	12	6	Good	Poor	.94	1.10
2	20/30	20/30	.....	.....	0	0	0	0	0	0	0	0	4	0	4	0	Poor	None	.76	1.00
3	20/30	20/40	20/25	20/30	0	0	4	6	6	4	12	2	14	2	0	0	Good	Poor	.97	1.90
4	20/40	20/40	20/25	20/30	0	2	4	0	9	4	12	8	12	8	17	8	Poor	None	.80	1.40
5	20/25	20/25	.....	.....	0	0	2	0	8	6	9	6	8	5	26	20	Good	Poor	.83	1.11
6	20/25	20/25	.....	.....	0	0	10	0	7	3	7	1	6	2	4	0	Good	Good	1.11	2.00
7	20/30	20/30	.....	.....	0	0	4	0	8	3	14	4	14	8	15	0	Poor	None	.76	1.10
8	20/20–3	20/20–4	.....	.....	0	2	0	4	8	5	18	6	12	0	22	14	Good	Poor	1.00	1.11
9	20/20–2	20/20–1	.....	.....	2	0	4	0	8	6	8	4	18	12	20	14	Poor	Poor	.88	1.40
10	20/25	20/25	.....	.....	2	0	8	0	14	6	6	0	16	14	8	0	Poor	None	.94	1.20
11	20/40	20/40	20/25	20/30	0	0	0	2	8	4	24	6	10	3	40	30	Poor	Poor	0.79	1.80

S: Without.
C: With.

are inclined to believe that certain children are not ready to read although they have reached their sixth birthday; and visual immaturity should be recognized as an important factor in this lack of reading readiness.

There are several aspects of visual immaturity. If all the children entering the first grade were tested for visual acuity it would be found that many do not have full 20/20 vision, while at the end of the second school year most of them do have normal vision. This may simply be due to the circumstance that the child has not yet learned to translate small retinal images into symbols with meanings. Such low visual maturity in children entering school has been reported by several observers. Berner and Berner[15], as well as Lancester,[13] observed this factor and discussed its importance in early development.

We found that of thirty-seven first-graders—59 per cent—had a visual acuity which was below 20/20; of the second-graders, 54 per cent; while with children of the third, fourth, fifth and higher grades the percentage of those with visual acuity below 20/20 was as low as 20 per cent and never higher than 35 per cent. Furthermore, in the older children low visual acuity was nearly always accompanied by a corrective refractive error, while of the eleven pre-readers only three needed glasses; and of a group of thirty-seven first-graders only seven out of twenty-two who had vision below 20/20 needed glasses. The other fifteen were low because of visual immaturity.

Table I gives the total records of eleven first-graders before they began to read. The reading score is one of reading readiness and not reading accomplishment. The mental age expectancy score indicates what a child should be able to do according to his mental age. Out of the whole group only one had 20/20 visual acuity although there are some who are only slightly below 20/20. Only four had good fusion, while the others had a tendency to suppress or alternate, especially at the near point. This weakness or immaturity in binocular vision is still more apparent in stereopsis. Only one had definitely established stereopsis, four had none, and the rest attained it for simple pictures only. Most of these children had low ductional reserve, while one child had too much. This child also had an uncorrected hypermetropic refractive error causing too much accommodation and convergence. It is also interesting to note that the children with weak

binocular vision all made lower scores on the reading-readiness test than would be expected from their mental age expectancy.

The data from these eleven youngest children, ranging from the age of 5.8 to 6.5 years, are presented as an example of the findings obtained for the entire group of the first- and second-graders. Visual maturity develops fast during the first two years of school. Lack of space prevents the demonstration of this fact from all the measurements obtained on all the children from these grades. However, there is presented a comparison of the percentage of incidence of low eye measurements in various grade levels.

TABLE II.—PERCENTAGE OF INCIDENCE OF VARIOUS EYE FACTORS, AT DIFFERENT GRADE LEVELS, INDICATING IMMATURITY AND ABNORMALITY

	Pre-readers	1st Grade	2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade
Total Subjects	11	26	28	10	11	10	18	23	89
Vision below 20/20	54%	59%	54%	20%	25%	20%	35%	27%	34%
Squint	0%	3%	4%	10%	0%	10%	6%	5%	0%
Good fusion	54	67	78	70	73	90	82	86	80
Poor Fusion	46	18	22	20	27	10	18	9	16
No Fusion	0	15	0	10	0	0	0	5	4
Good Stereopsis	18	29	57	50	64	80	71	82	75
Poor Stereopsis	27	25	33	30	35	20	18	13	14
No Stereopsis	55	45	10	20	10	0	11	5	11
Distance Phorias	36	13	7	10	29	10	0	23	10
Near Phorias	63	52	32	60	54	40	39	64	46
Poor ductions (distance) — Abnormal Abduction break	9	10	9	20	9	10	16	5	4
Poor ductions (distance) — Abnormal Abduction Return	27	13	14	30	9	10	5	9	10
Poor ductions (distance) — Abnormal Adduction Break	72	55	35	40	63	40	16	36	40
Poor ductions (distance) — Abnormal Adduction Return	81	73	50	50	32	50	51	56	57
Poor ductions (near) — Abnormal Abduction Break	63	55	61	20	36	20	16	18	25
Poor ductions (near) — Abnormal Abduction Return	72	68	64	45	52	40	44	27	44
Poor ductions (near) — Abnormal Adduction Break	72	63	45	50	54	30	11	35	25
Poor ductions (near) — Abnormal Adduction Return	90	70	56	50	48	50	27	34	46

In the first and second grades there are more children who have visual acuity less than 20/20 than in any of the other grades.

There are fewer children with good fusion and stereopsis in these two lowest grades. Also more children have weak ductions, particularly poor adduction and duction reserve, manifested by the measurement of 'fusion return.' 'Fusion return' is the point where diplopia disappears when the prism strength is diminished after having appeared when the strength of the rotary prism was increased until the limit of fusional amplitude had been reached. This adduction weakness, although present for distant vision, is especially marked at the near point. This fusional convergence deficiency results in occasional monocular suppression, the left eye usually being the one suppressed by the right-handed child and the right eye usually being suppressed by the left-handed child. Reading becomes slow and difficult because the dominant eye reads the words first while the other eye causes confusion. This is particularly true when the child is reading long words or sentences. During the time necessary for interpretation of the words, binocular vision is temporarily restored and the print seems to change in appearance and position. Children frequently tell us that words jump or the letters become confused. Others will volunteer the information that upon looking at a word the rest of the page becomes black, which necessitates rereading of the sentence. This condition, if it continues unchecked, may predispose to reading difficulty.

If these conditions are found to be prevalent in eyes that are not functionally mature, and if such conditions may cause reading difficulty, it is very important that normal ocular habits be established as early as possible. The child should be given the necessary time for his visual mechanism to develop so that reading becomes possible, or, if ocular abnormalities become established and tend to persist, binocular vision should be taught and improved by proper exercises. For if a child is forced to learn to read before good binocular integration is developed, ocular difficulties probably tend to become fixed and develop into such abnormal conditions as phorias, squint, poor convergence, poor fusion and abnormal fusional reserve, and poor stereopsis. Such children not only may have difficulty in learning to read, but, when urged to do much close work, may develop nervousness and irritability. The anxiety aroused in them as a result of criticism from teachers and parents, and their own knowledge of

their failure to learn to read, may easily cause psychological behavior problems.

If a child has not learned to read and ocular difficulties are present, it is wise to correct these difficulties with proper orthoptic training and to repeat the steps used in the early teaching of reading. However, much of this corrective work might be avoided by proper supervision and guidance in the development of binocular habits during the first and second grades. Much trouble might be saved if the child could receive a thorough eye examination including muscle tests and tests of binocular integration when the actual teaching of reading is begun if at that time symptoms of reading difficulties appear.

In conclusion it should be restated briefly that many of the first- and second-graders who were tested showed a visual mechanism which was incompletely matured. This was evident among these children by the frequency of low visual acuity, poor fusion and stereopsis, low and unstable ductions, and fusion amplitude as compared with the children from the other grade levels. These abnormal eye conditions were particularly noticeable for the near point. Since abnormal eye conditions, and especially abnormal binocular integration, may either cause or contribute to reading difficulties at whatever age they may occur, it is of prime importance that a child be free from such handicaps at the time he has to begin the complicated process of learning to read. Careful checks and visual guidance of those with reading difficulties during the first grade, or before starting school, might prevent many serious reading problems as well as abnormal eye conditions.

BIBLIOGRAPHY

1) Park, G. E., and Burri, C., "Visual Factors in Reading Difficulty," *J. Ed. Psychol.*, May, 1943.

2) Park, G. E., and Burri, C., "The Effect of Various Eye Conditions on Reading Difficulty," *J. Ed. Psychol.*, October, 1943.

3) Burri, C., "Process of Learning Simultaneous Binocular Vision," *Arch. of Ophthal.*, 1942, 28, 235–244.

4) Park, G. E., "An Investigation of the Angular Relation of the Visual (Visierlinie) and Optic (Corneal) Axes of the Eye," *Am. J. Ophthal.*, Nov., 1936, 19, 967–974.

5) Park, G. E., and Park, Russell Smith, "Further Evidence of Change in Position of the Eyeball During Fixation," *Arch. of Ophthal.*, June, 1940, 23, 1216–1230.

6) Preyer, W., *The Mind of the Child*, New York, D. Appleton, 1901.

7) Blanton, M. G., "The Behavior of the Human Infant During the First Thirty Days of Life," *Psychol. Rev.*, 1917, 24.

8) Shinn, M. W., "Notes on the Development of a Child," *University of California Studies*, 1893–09. Vol. I, 1–4.

9) Jones, M. C., "The Development of Early Behavior Patterns in Young Children," *Ped. Sem.*, 1926, 33, 537–585.

10) Ling, B. C. J., "A Genetic Study of Sustained Visual Fixation and Associated Behavior in the Human Infant from Birth to Six Months," *J. Genet. Psychol.*, 1942, 61, 227–277.

11) Johnson, B. and Bech, L. F., "The Development of Space Perception: 1. Stereoptic Vision in Pre-school Children," *J. Genet. Psychol.* 1941, 58, 247–254.

12) Mann, I. C., "Convergence Deficiency," *Brit. J. Ophthal.*, 1940, 24, 373.

13) Lancester, W. B., "Fifty Years Experience with Ocular Motility," *Amer. J. Ophthal.*, 1941, 24, 741–748.

14) Nugent, D. G., and Ilg, V., "Newer Developments in Orthoptics with Reference to Reading Problems," *Arch. Phys. Therapy*, 1941, 22, 225–232.

15) Berner, G. E. and Berner, D. E., "Reading Difficulties in Children," *Arch. of Ophthal.*, 1938, 20, 829–838.

EYE MOVEMENTS IN READING TYPE SIZES IN OPTIMAL LINE WIDTHS*

DONALD G. PATERSON AND MILES A. TINKER

University of Minnesota

In studying the effects of type size on readability it is necessary to take account of variations in line width and leading. The writers have reported the influence of varying amounts of line widths and leading on speed of reading for each of the more commonly used type sizes.[1] On the basis of the obtained results it was possible to designate optimal line widths and leading for each of six type sizes. A comparison of these optimal arrangement revealed that both eight-point and six-point type were read significantly slower than the larger sizes (nine, ten, eleven and twelve point) which were read with equal speed.[2]

In order to provide an analytical picture of the oculomotor patterns accompanying the retardation in reading each of the smaller type sizes in comparison with one of the larger ones (eleven point), the present investigation was undertaken. The Minnesota eye-movement camera was employed to photograph the eye movements.

In the first comparison, each of twenty college students read ten paragraphs from the Chapman-Cook Speed of Reading Test, Form A, set in eleven-point Granjon type with optimum line width (twenty-two picas) and leading (two points), and ten different paragraphs from Form B of the same test in eight-point Granjon type with optimum line width (sixteen picas) and leading (two points). In the second comparison, each of twenty additional college students read ten paragraphs from Form A in the eleven-point Granjon (twenty-two-pica line width, two-point leading) and ten paragraphs from Form B in six-point Granjon with optimum line width (fourteen picas) and leading (two points). Figure 1 shows samples of the reading material used in both studies.

* The writers are grateful to the University of Minnesota Graduate School for research grant to finance this study.

[1] D. G. Paterson and M. A. Tinker, *How to Make Type Readable.* New York: Harper and Bros., 1940, pp. 209.

[2] D. G. Paterson and M. A. Tinker, *op. cit.*, pp. 80–81. These data were obtained in 1941 by Howard Roy in a study conducted for his M. A. thesis at University of Minnesota.

ELEVEN-POINT VERSUS EIGHT-POINT TYPE

The results for the comparison between eleven-point and eight-point type are given in Table 1. It will be noted that, in comparison with the eleven-point text, material in eight-point was read with more fixations, fewer words per fixation, a longer pause duration, a longer perception time and with more regressions. With the exception of regression frequency, all the

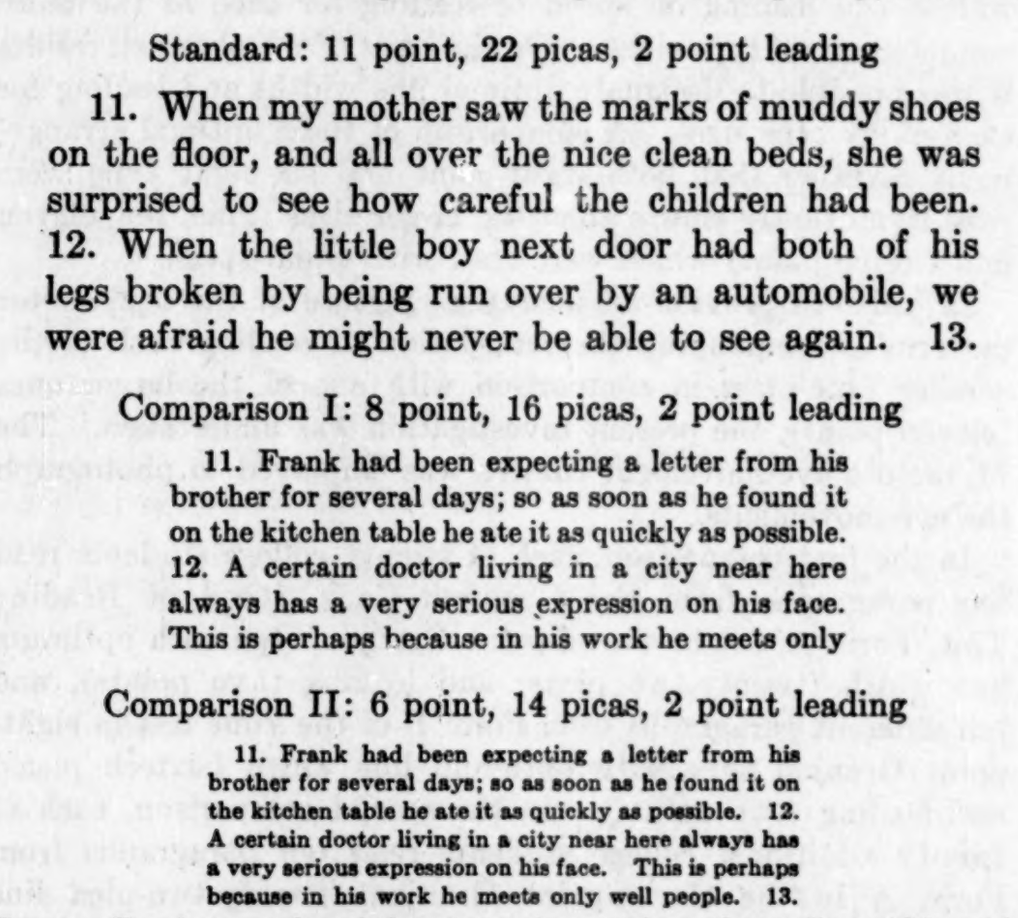

Standard: 11 point, 22 picas, 2 point leading
11. When my mother saw the marks of muddy shoes on the floor, and all over the nice clean beds, she was surprised to see how careful the children had been. 12. When the little boy next door had both of his legs broken by being run over by an automobile, we were afraid he might never be able to see again. 13.
Comparison I: 8 point, 16 picas, 2 point leading
11. Frank had been expecting a letter from his brother for several days; so as soon as he found it on the kitchen table he ate it as quickly as possible. 12. A certain doctor living in a city near here always has a very serious expression on his face. This is perhaps because in his work he meets only
Comparison II: 6 point, 14 picas, 2 point leading
11. Frank had been expecting a letter from his brother for several days; so as soon as he found it on the kitchen table he ate it as quickly as possible. 12. A certain doctor living in a city near here always has a very serious expression on his face. This is perhaps because in his work he meets only well people. 13.

FIG. 1.—Showing samples of three type sizes set in optimum line widths and leading.

differences are significant beyond the one per cent or the five per cent level. In other words, the slower rate of reading the eight-point type is reflected both in more fixations and longer duration of the fixational pauses. Perception time is merely a combination of fixation frequency with pause duration. Since the number of words read in Forms A and B was constant, the decrease in words per fixation for the eight-point type reflects the increase in fixation frequency. The lack of significance of

a 7.8 per cent difference for regression frequency is probably due to its large relative variability in comparison to the other measures. It will be noted, however, that the difference is in the same direction as the others.

TABLE 1.—MEAN EYE-MOVEMENT MEASURES: ELEVEN-POINT *vs.* EIGHT-POINT TYPE

Note: Pause duration and perception time are reported in seconds. Ten paragraphs from the Chapman-Cook Speed of Reading Test, Form A, were set in eleven-point type with a twenty-two pica line width and ten paragraphs from Form B of the same test were set in eight-point type with a sixteen-pica line width. All paragraphs were printed as follows: Granjon, lower case, two-point leading, on eggshell paper stock. N = 20 university students.

Type Size	Fixation Frequency	Words per Fixation	Pause Duration	Perception Time	Regression Frequency
11 point.....	186.3	1.7	0.22	41.7	28.7
8 point.....	198.8	1.5	0.23	46.1	31.0
Differences*.....	+12.5	−0.2	+0.01	+4.4	+2.3
Per cent Diff.....	+6.7	−7.8	+4.2	+10.5	+7.8

* The *t* values for pause duration and perception time are beyond the one per cent level; for fixation frequency and words per fixation beyond the five per cent level; for regression frequency beyond the twenty per cent level.

ELEVEN-POINT VERSUS SIX-POINT TYPE

The results for the reading of eleven-point and six-point type are given in Table 2. Again the differences are all in the same direction, indicating less efficient reading of six-point type. The six-point type was read with more fixations, fewer words per fixation, a greater pause duration, a greater perception time and more regressions. Differences for pause duration and perception time, however, are the only ones which are significant beyond the one per cent level. These differences amount to 7.9 and 9.9 per cent, respectively. Thus, when set in optimal line widths and leading, the fundamental difference in oculomotor patterns for reading six-point and eleven-point type is that fixational pauses have longer duration on the average. The

increased perception time again reflects the combination of pause duration with fixation frequency.

TABLE 2.—MEAN EYE-MOVEMENT MEASURES: ELEVEN-POINT *vs.* SIX-POINT TYPE

Note: Pause duration and perception time are reported in seconds. Ten paragraphs from the Chapman-Cook Speed of Reading Test, Form A, were set in eleven-point type with a twenty-two-pica line width and ten paragraphs from Form B of the same test were set in six-point type with a fourteen-pica line width. All paragraphs were printed as follows: Granjon, lower case, two-point leading, on eggshell paper stock. N = 20 university students.

Type Size	Fixation Frequency	Words per Fixation	Pause Duration	Perception Time	Regression Frequency
11 point.....	194.4	1.62	0.24	46.6	26.2
6 point.....	197.8	1.59	0.26	51.2	27.9
Differences*.....	+3.4	−0.03	+0.02	+4.6	+1.7
Per cent Diff.....	+1.8	−1.9	+7.9	+9.9	+6.7

* The *t* values for pause duration and for perception time are significant beyond the one per cent level; for fixation and regression frequency, beyond the fifty per cent level; for words per fixation, beyond the forty per cent level.

The eye-movement patterns revealed in reading the six-point and the eight-point type are essentially alike. For both sizes of type the eye-movement measures indicate uniformly less efficient reading. While word forms are not noticeably distorted in the smaller type sizes, visibility of the printed words is considerably less in comparison with the eleven-point type.[1] Presumably, this reduction in visibility necessitates longer pause durations and increased fixation frequency.

SUMMARY

1) In a series of reading performance experiments the writers determined optimal line widths and leading for each of the commonly used type sizes. These studies of type sizes set in

[1] M. Luckiesh and F. K. Moss, *Reading as a Visual Task.* New York: D. Van Nostrand Co., 1942, pp. 128–130.

these optimum line widths and leading showed that six- and eight-point type were read more slowly than larger type sizes.

2) Eye-movement photographs of two new groups of twenty subjects each indicate the precise changes in oculomotor patterns involved in reading six- and eight-point type in comparison with eleven-point type.

3) In reading eight-point type there was a significant increase in fixation frequency, pause duration and perception time and a significant decrease in words read per fixation. Regressions were also slightly increased.

4) For six-point type there was a significant increase in pause duration and perception time. The other measures, although not significantly different, were in line with the findings for eight-point type.

5) Reduced visibility of the smaller type sizes appears to be the main factor producing the changes in oculomotor behavior.

PART AND WHOLE METHODS IN THE MEMORIZATION OF MUSIC

CYRIL C. O'BRIEN

Department of Psychology, Maritime Academy of Music, Halifax, N.S., Canada

The factor of the economy of time in the memorization of music has frequently been the object of much speculation. Is it better to learn a selection of music—studying it from beginning to end—or memorize it by concentrating upon one part at a time? Such a question must necessarily take into consideration the numerous types of instruments in addition to vocal memorizing, but even amid this diversified array, three kinds of memory stand forth; namely, visual, auditory and tactile. The memorization of a piano selection would involve these three types of memories, and any study into the relative economy of time in the utilization of part and whole methods must take cognizance of the three factors.

In the investigation of this problem, four graduate students of the Maritime Academy of Music, Halifax, N.S., acted as subjects. Experiments were conducted by comparing part and whole methods of memorizing music after an isolation of factors comprising composite memory.* All experiments were performed with the aid of metronome and stop-watch. Each selection memorized was well within the range of technique of the performers concerned.

Two piano selections were given for Experiment 1. The first was "Valse" from *Le Petit Carnaval* by L. Streabbog, Opus 105, No. 1—a waltz of thirty-two measures in the key of G: the other, purposely named "Valse Plagiarist", was patterned on the first. This second composition also contained thirty-two measures, was written in the same key and had exactly the same number and kind of notes and rests, although sufficiently different in melody and contour to mitigate any resultant practice effect.

To ensure that Selections No. 1 and No. 2 were equal in degree of difficulty before proceeding with the experiment, ten members of the Piano Department Staff of the Maritime Academy of Music were asked to grade each selection. In determining the

* There was also a Gestalt approach to the problem by using three senses combined—auditory, visual and tactile—as, in Experiments 1 and 2.

grade to be stipulated, each piano teacher reached his decision without consulting other members of the Staff. The results of this brief survey are tabulated in Table I.

TABLE I.—GRADING BY PIANO TEACHERS OF THE TWO SELECTIONS

Teacher	Grade of Selection 1	Grade of Selection 2
No. 1	1½	1½
No. 2	1	1
No. 3	1½	1½
No. 4	1½	1½
No. 5	1	1
No. 6	1½	1½
No. 7	1	1
No. 8	1	1
No. 9	1	1
No. 10	2	2

As will be noted in the table, all ten teachers graded the selections as having the same degree of difficulty: that is, the relative value assigned by each was the same. Furthermore, as stated previously, in number and kind of notes and rests, the selections were similar. Mathematically, each piece had the identical number of notes and rests.

Selection No. 1 was learned by the whole method. The metronome was set at 120. Definite and uniform instructions were given to place score upon piano, set stop-watch going and begin playing immediately from beginning to end. When the subject believed the selection to be known, the watch was stopped, (always at the end of a selection), and the time recorded in seconds. Then the composition was played from memory once. If even one note was incorrect, the stop-watch was again set going, the selection being played from beginning to end until the subject once more believed it to be known. The watch was stopped at once and the time set down. This procedure was continued until the selection could be played accurately from memory. The time of each trial was recorded in seconds and the total time for all trials obtained.

Selection No. 2—'Valse Plagiarist' built upon the 'Valse' by Streabbog—was memorized according to the previous directions, with the exception of being learned by the part method, four

measures being memorized at a time. The metronome was set at 120. The total time was obtained as in the whole method. The experiments were timed one to two weeks apart. Uniformity of directions and attack were applied for the memorization of each composition. The subjects were given the selections in rotation in order to eliminate experimentally practice effect.

"Valse" from *Le Petit Carnaval* and "Valse Plagiarist" each contains thirty-two measures, two hundred five notes and sixteen rests. The subjects who carried out the various trials are designated—A, B, C, D. The results of the first experiment are recorded in Tables II, III, and IV.

TABLE II.—TIME IN SECONDS FOR LEARNING SELECTIONS

Pupil	Part Method	Whole Method
A	68.2	195.
B	163.0	244.
C	505.0	720.
D	482.4	660.

TABLE III.—SECONDS PER NOTE BY PART AND WHOLE METHOD

Pupil	Part Method	Whole Method
A	.332	.951
B	.795	1.190
C	2.463	3.512
D	2.353	3.219

TABLE IV.—TIME SAVED PER NOTE (IN SECONDS) AND PER CENT OF TIME SAVED USING PART METHOD

Pupil	Time Saved	Per Cent
A	.619	65.08
B	.395	33.12
C	1.049	29.86
D	.866	26.90

The pupils had no preference in the matter of the two selections. According to their written statement, there was no predilection for either composition.

The two selections used in Experiment 2 were "Toy Soldiers" by Ada Richter (Whole Method) and "Peter Pump" by David Slayter (Part Method). Both compositions are in four-four time. "Toy Soldiers" contains one hundred ninety-four notes and "Peter Pump" three hundred eighty notes. Since the

selections used in Experiment 1 showed the part learning to advantage, "Peter Pump," containing nearly twice the number of notes as "Toy Soldiers," was chosen as the march to be memorized by the part method—four measures at a time. The metronome was set for the memorization of each selection at 120. The rules and directions for Experiment 2 were the same as in Experiment 1. The piano teachers rated the two music pieces as shown in Table V.

TABLE V.—RATING BY PIANO TEACHERS ON COMPOSITIONS USED IN EXPERIMENT 2

Teacher	Grade of "Toy Soldiers"	Grade of "Peter Pump"
No. 1	2	2½
No. 2	1½	2½
No. 3	1½	2½
No. 4	1½	2½
No. 5	1½	2½
No. 6	1	2
No. 7	1	1½
No. 8	1½	2½
No. 9	½	2½
No. 10	1½	2½
	13.5	23.5

TABLE V*A*.—RESULTS IN LEARNING COMPLETE COMPOSITIONS ENUMERATED IN SECONDS

Pupil	Part Method	Whole Method
A	308.	220.6
B	209.8	
C	196.	330.
D	559.4	363.8

Taking the average grade according to Table V, "Toy Soldiers" is 1.35, while "Peter Pump" is 2.35—exactly one grade in advance. In matter of choice, Pupil A preferred "Toy Soldiers," and Pupils B, C and D liked "Peter Pump" better.

When the auditory, visual and tactile phases of memory were employed in the pianistic memorizing of the foregoing experiments, there was a decided saving of time in the utilization of the part method.

TABLE VI.—SECONDS PER NOTE BY PART AND WHOLE METHODS

Pupil	Part Method	Whole Method
A	.810	1.137
B	.552	
C	.515	1.701
D	1.472	1.875

TABLE VII.—TIME SAVED PER NOTE AND PER CENT OF TIME SAVED USING PART METHOD

Pupil	Time Saved	Per Cent
A	.327	28.75
B		
C	1.186	69.73
D	.403	21.38

SOUNDLESS MEMORIZING

Experiment 3 comprised in the main a consideration of soundless memorizing. "Lola Waltz," Key of C, by Emil Ascher (Whole Method) and "Lilac Waltz," Key of F, by Emil Ascher (Part Method) were the selections used. Both pieces contain thirty-two measures. Directions were given to read "Lola Waltz" from beginning to end, associating the letter name with each note. Date and time of experiment were set down. The selection was not sung at sight. The notes were visualized. The letter name was associated with the note, but not the pitch. The time and rhythm were studied by memorizing the note value of each note. The metronome was set at 120. As soon as the stop-watch was set going, "Lola Waltz" was read from beginning to end. When the subject thought that the melody could be written unaided, the watch was stopped, time marked down and the thirty-two measures written as well as possible. Comparison was then made with the original. If any mistakes had been made, they were noted, the stop-watch started and the selection read from beginning to end again. When thought to be known, the watch was stopped and time set down as previously. The thirty-two measures were again written on manuscript paper, care being taken that the first copy was not seen. The above procedure was repeated until the thirty-two measures could be written without any mistakes. All time was then totalled. The object of the experiment was to eliminate as much as possible the factor of sound.

"Lilac Waltz," the second selection in this experiment, was read four measures at a time, the watch being stopped when it was considered that the measures were known. Time was recorded for every four measures. When this sectional memorizing was complete, all thirty-two measures were written down. If any connection between two groups or a part of a group could not be recalled, the watch was started and the part learned. All times were totalled. In the memorization of this melody by the part method, the pitch factor was eliminated as in learning the previous one by the whole method.

TABLE VIII.—RESULTS OF THE SOUNDLESS MEMORIZING (IN SECONDS)

Pupil	Part Method	Whole Method
A	271.0	356.66
B	202.2	429.00
C	121.0	615.00
D	293.4	581.60

TABLE IX.—PER CENT OF TIME SAVED USING PART METHOD

Pupil	Per Cent
A	23.42
B	52.90
C	80.32
D	49.55

AUDITORY MEMORIZING

Experiment 4 was termed Auditory Memorizing. The first sixteen measures of "Little Sweetheart" by H. P. Hopkins was used for the experiment in Whole Method and a selection patterned on this, different in melody, but equated in kind and number of notes, was used for the Part Method. In the part method, four measures were played on the piano and the subject instructed to write the notes upon manuscript, the key having been given. The entire sixteen measures were played in this manner three times. Pitch and time errors were noted. In the whole method, the sixteen measures were played from beginning to end three times. The subjects were instructed to write the notes upon the manuscript paper provided for the purpose. The key was given as previously and the metronome marking in each

case was 104. The results showing the number of pitch and time errors for each method are recorded in Table X.

TABLE X.—NUMBER OF PITCH AND TIME ERRORS—EXPERIMENT 4

Pupil	Part Method		Whole Method	
	Time	Pitch	Time	Pitch
A	0	0	0	0
B	0	0	6	5
C	0	2	8	6
D	0	4	0	2

In Experiment No. 5 the melody and words of two songs—equal in number and kind of notes—were memorized, one by part and the other by whole method. "In Heavenly Love Abiding" was memorized by part method and "Roses" by the whole method. The first sixteen measures of the latter selection were learned in this manner. It was first thought, since the whole method in many instances favours the memorization of poetry, that the combination of the words with the melody would yield positive results for the whole method. However, the figures in Table XI indicate that the returns are neither positive nor negative:

TABLE XI.—MEMORIZING OF MELODY AND WORDS (RECORDED IN SECONDS)

Pupil	Part Method	Whole Method
A	134.0	245.2
B	292.0	93.0
C	160.0	152.0
D	183.4	304.0

In Experiment No. 6 a pipe organ minus the wind (with current turned off) was used. The selection "Pretty Rosebuds" by H. P. Hopkins, Opus 146, No. 1, was learned by the whole method as far as the end of the sixteenth measure. For the part method, another selection different in construction and design but equal in number and type of notes was composed similar to the former. Directions were like those stated in Experiment

No. 1, with the exception of the factor of sound being eliminated. The metronome was set at 120. The kinaesthetic and visual capacities predominated in this test. Table XII shows the results of the experiment.

TABLE XII.—RESULTS (IN SECONDS) OF LEARNING IN EXPERIMENT 6

Pupil	Part Method	Whole	Per Cent of Time Saved Using Part Method
A	27.6	65.4	57.79
B	155.0	250.0	38.00
C	32.4	90.0	64.00
D	110.0	168.6	34.75

The research reported here and which was carried out at the Maritime Academy of Music in Halifax, required nine months. At least one week intervened between two successive sections of a part and whole experiment and of times two weeks passed before the concluding phase of an individual experiment was completed. The resultant effect of practice was reduced to a minimum in every case.

Three of the four subjects took both forms of the Drake Musical Memory Test and all wrote the complete batteries of the Seashore and the Kwalwasser-Dykema Tests. Table XIII gives the norms for the Drake Musical Memory Test together with those of the Seashore and K-D Tests for Tonal Memory.

TABLE XIII.—NORMS OF THE FOUR SUBJECTS

Pupil	Drake	Seashore	K-D
A	98	97	95
B	88	75	84
C	..	79	84
D	98	75	98

SUMMARY

1) In the consideration of all results and of the subjects studied, the part method of the memorization of music was more economical.

2) For the piano selections used in which auditory, visual and tactile memories were employed, there was definite saving of time ranging from twenty-five per cent to sixty-five per cent in the utilization of the part method.

3) In the soundless memorizing of music (visual memory alone), the part method was definitely superior in economy of time.

4) Neither whole nor part method was favoured when all the results of the experiment in auditory memorization were considered.

5) In the learning of melody and words, the figures are partly positive and partly negative.

6) Where the kinaesthetic memory was predominant, tactile memorization of music by the part method was accomplished with a saving of time ranging from thirty-five per cent to sixty-five per cent.

Perhaps there are individual idiosyncrasies in the learning of music, so that certain individuals learn most rapidly by the part method and others by the whole method. There is some evidence in favour of such an assumption. Accordingly, a second assumption might be made that the frequency of this type of individual in the general population is fifty-fifty. Therefore, the likelihood that the four pupils all prefer the part method of learning is $(\frac{1}{2})^4$ or one in sixteen—not an unlikely occurrence. Taking and interpreting the data of this study, however, it appears that there is evidence that in the memorization of music involving auditory, visual and tactile capacities, when the music is well within the range of the technique of the subjects the part method of memorizing is more economical.

THE EFFECT OF CONTEXT ON LEARNING A VOCABULARY

C. L. MORGAN AND WILBERT L. BAILEY

Purdue University

PROBLEM AND PROCEDURE

The problem of this study consisted of an evaluation of the effects of context upon: (a) the recognition of English equivalents of an artificial vocabulary, (b) the time required for the completion of the learning exercises, and (c) the number of words looked up in the dictionary. Context is used here in the sense of connected meaning.

The learning materials consisted of a short story exercise; a second exercise identical in general form and content except for the absence of the story; a dictionary; and a test. Excerpts from the French story "La Tour Eiffel"[1] were used for the short story exercise. To avoid familiarity with the foreign words employed in this story, an artificial vocabulary, which consisted of one hundred and fourteen words, was substituted. A second exercise which was an exact duplicate of the first in general form and content, except for the absence of the story, was constructed. By changing the word sequence of the contexual material, while retaining the same word frequencies and frequency sequences, the context was destroyed without otherwise affecting any other aspect of the learning problem.

A dictionary giving the English equivalents of the foreign words was prepared. All grammatical constructions were given in the dictionary. Columns for recording the number of times each word was looked up were provided directly opposite the meaning of the words. The vocabulary test contained all of the artificial words employed, with blank spaces provided in which to write the English equivalents. The tests were identical for both groups.

The population used in the experiment was selected from a sophomore class in Education Psychology at Purdue University. Two groups of forty-two students each were selected and matched on the basis of the percentile scores on the American Council

[1] Grace Cochran, and Helen M. Eddy, *Si Nous Lisions*, University of Chicago Press, 1929, pp. 18–24.

Psychological Examination and the Iowa Silent Reading Test. The distribution of the percentile scores for the two groups ranged from the ninety-seventh to the first percentile. Each group was composed of twelve subjects in the upper fourth, nineteen in the middle group, and eleven in the first fourth.

Because of conflicts in class schedules, it was impossible to have all the students of both groups meet on the same day of the week. A working schedule was arranged in which students of the two groups met once each week for five consecutive weeks. At no time was there a lapse of more than seven days between these meetings. For the purpose of motivating the groups, they were told that they would be excused from their final examination in Educational Psychology if and when they had completed the requirements of the experiment satisfactorily.

The Control Group was given the printed matter without context. The Experimental Group was given the exercise containing the story derived from "La Tour Eiffel." The first period was given over to carefully instructing both groups as to the procedure and requirements of the experiment. The groups were told that this was an exercise in learning a foreign vocabulary, a vocabulary that had been constructed for the purpose of this experiment. They were required to make a translation of the exercises, but were not permitted to write the translation. They were told that in order to complete the translation, it would be necessary for them to use their dictionaries; and a demonstration was given showing how to record the number of times each word was looked up in the dictionary. As a signal that the translation was completed, the subjects were directed to raise their hand and the time required for the translation was recorded. Records were secured for the groups for the time required for the completion of the learning exercises and the number of times words were looked up in the dictionary.

The tests were administered at the beginning of each learning period, with the exception of the first, and immediately following each learning period. Thus there was obtained an immediate and a delayed vocabulary test score for each person, for each of the five weeks during which the experiment was conducted. In accordance with the directions given to the subjects, no synonyms other than those given in the dictionary were allowed on these tests and there were no penalties for misspelling. The

directions concerning test performance made for complete objectivity of scoring. All papers were collected at the completion of each meeting. The subjects were not told that one group had a story to translate while the other did not. No inquiries were made of the experimenter as to the connection of the work of the two groups.

RESULTS

The correlation between the matched scores for the two groups on the Psychological Examination and on the Iowa Silent Reading Test was found to be .998 and .990, respectively. The correlation between these tests for the combined groups was found to be .969. These correlations, it will be observed, were spuriously produced because of the manner in which the population was selected and, therefore, are indicative only of the degree of care employed in matching the groups.

It will be observed from Table 1 that the differences in median scores for immediate recall are statistically unreliable, except for the fourth week where the difference in favor of the control group was statistically significant. The trend of obtained differences from week to week, as shown by the critical ratios, suggest a possible interference of the context in learning vocabulary. This hint, however, can only be verified by extended research.

TABLE 1.—COMPARISON OF GROUPS FOR IMMEDIATE RECALL ON VOCABULARY TESTS

Week	Control (Non-Context)		Experimental (Context)		Difference	PE of* the Diff.	Critical Ratio
	M'd'n Score	PE Med'n	M'd'n Score	PE Med'n			
1	40.75	2.50	39.80	1.93	.95	3.06	.31
2	72.00	2.29	64.10	2.52	7.90	3.35	2.36
3	89.00	2.48	78.60	2.06	10.40	3.16	3.29
4	98.30	1.01	87.50	2.34	10.80	2.54	4.25
5	106.00	1.49	97.30	2.33	8.70	2.76	3.15

* Formula for matched groups used throughout this study.

Table 2 shows no differences which are statistically significant. There is even less indication of any suggestive trend for delayed recall than was true of immediate recall. The obtained differences are consistently in favor of the control group, as was true of immediate recall data, and this again is suggestive only.

TABLE 2.—COMPARISON OF GROUPS FOR DELAYED RECALL ON VOCABULARY TESTS

Week	Control (Non-Context)		Experimental (Context)		Difference	PE of the Diff.	Critical Ratio	In favor of
	M'd'n Score	PE Med'n	M'd'n Score	PE Med'n				
1	23.00	1.44	19.80	1.07	3.20	1.78	1.80	Control
2	50.50	2.29	44.20	1.81	6.30	2.86	2.20	Control
3	70.70	2.27	62.50	2.09	8.20	3.06	2.68	Control
4	85.00	2.47	77.00	1.86	8.00	3.07	2.58	Control
5	96.90	2.64	86.50	1.74	10.40	3.03	3.43	Control

TABLE 3.—COMPARISON OF GROUPS ON WEEKLY TIME SCORES IN MINUTES TAKEN FOR THE TRANSLATION

Week	Control (Non-Context)		Experimental (Context)		Difference	PE of the Diff.	Critical Ratio
	M'd'n Score	PE Med'n	M'd'n Score	PE Med'n			
1	66.28	2.30	60.70	1.98	5.58	2.83	1.97
2	46.93	1.80	41.50	2.39	5.43	2.64	2.06
3	36.50	1.45	30.75	2.01	5.75	2.32	2.48
4	31.50	1.57	26.00	1.59	5.50	1.94	2.83
5	20.90	1.64	20.64	1.33	.26	2.02	.13

Table 3 presents no differences in time taken which can be accepted as reliable. The differences obtained, however, are consistently in favor of the experimental group. The same pic-

ture is true for Table 4 which represented the frequency of use of the dictionary.

TABLE 4.—COMPARISON OF GROUPS ON THE NUMBER OF TIMES WORDS WERE LOOKED UP IN THE DICTIONARY

Week	Control (Non-Context)		Experimental (Context)		Difference	PE of the Diff.	Critical Ratio
	M'd'n Score	PE Med'n	M'd'n Score	PE Med'n			
1	220.50	5.51	194.50	6.67	26.00	8.50	3.06
2	147.50	7.79	106.00	7.38	41.50	10.58	3.92
3	77.17	6.35	59.50	4.38	17.67	7.57	2.33
4	49.50	4.70	34.38	3.66	15.12	5.84	2.59
5	29.83	4.22	23.00	2.62	6.83	4.83	1.41

The results make it impossible to draw any conclusions, except the general fact that presence or absence of context did not make for statistically significant difference in use of dictionary, time taken in translation, delayed recall or immediate recall, except in the latter instance for the fourth week.

The results do suggest that with a more adequate sampling and extended materials, context may bring less frequent use of dictionary, reduce the time spent and, as a total result of these two factors, reduce the amount of learning as measured by vocabulary recall tests. It should possibly be further suggested that presence of context may lead the learner to guess the meaning of words and not verify his guesses.

BOOK REVIEWS

MARGARET LOUISE HAYES. *A Study of the Classroom Disturbances of Eighth-grade Boys and Girls.* New York: Bureau of Publications, Teachers College, Columbia University, 1943, pp. 139.

What is the nature of disturbing behavior in eighth-grade boys and girls? What is the relationship of these behavior disturbances to selected factors in personalities and backgrounds of children and their teachers? These in brief are the chief problems investigated in *A Study of the Classroom Disturbances of Eighth-grade Boys and Girls.* Subjects for the study were sixty-eight children in the eighth grade of the Milne School in Albany, New York, and twenty-three practice-teachers. The investigators were aware of the fact that the classroom situation and the children investigated may not be an adequate sampling of eighth-grade children in general. The methods for investigation were composed of two types of observational techniques: (1) a time-sampling code method and (2) the recording of episodes of disturbing behavior. These materials were supplemented by stories written by the children, teacher interviews and letters, and records available in the institution concerning the backgrounds and personalities of pupils and teachers.

To a large extent the findings are in accord with previous studies. Illustrative of the findings are: Teachers reported more of the overt, noisy types of disturbances than observers recorded. The most frequent response of pupils to teachers' attempts to control their disturbing behavior was compliance although further disturbances occurred about half as often as compliance. Disturbing behavior seems to decrease with an increase in the number of pupils' interests, up to six or more interests. Teachers who themselves engaged in three or more leadership activities in college had fewer disturbances in their pupils than those who engaged in one. No significant relationships were found between disturbances by pupils and sex, intelligence, scholastic achievement in college, size, and general health. The chief causes of disturbing behavior given by the pupils were: wrong treatment by parent, previous frustrations, wrong companions, poor homes, a naturally bad nature, and dislike or hatred of school. The chief causes given by the teachers in order of frequency were:

the nature of pupils at the early adolescent age level, fatigue of teacher, lack of child participation, imitation of others, emotional tie with the recipient of the disturbance, directions not clear, and material not interesting. Some case study material of both children and teachers is presented but, all in all, the investigation is dominantly on the segmented rather than integrated level. Materials presented should be of some value to schoolteachers as well as to students of child behavior. H. MELTZER.

Psychological Service Center, St. Louis, Missouri.

CLARK L. HULL. *Principles of Behavior.* New York: D. Appleton-Century Co., 1943, pp. 422.

The task which the author has provided for himself is without question the most difficult which nature has to offer—"to present in an objective, systematic manner the primary, or fundamental, molar principles of behavior." In our present state of knowledge it is obvious that we must be satisfied with treating the relationships of general observable behavior (molar) rather than with the microscopic analysis of behavior in its neurophysiological interactions (molecular).

Hull's general procedure for developing the basic principles of behavior is to build up a number of 'logical constructs' which are intermediate to the stimulus and the response of organisms. Because these constructs are always tied to objectively measureable antecedent and consequent conditions, the probability is increased that the fundamental postulates and theorems of the system are not merely academic philosophizing.

The foundation upon which the hierarchy of principles is built is that organisms are adaptive in nature because of the screening processes of evolution. Adaptive reactions are based on muscular and glandular reactions in certain combinations. Some of these reactions are regular (inherent) while others tend to remedy the inadequacies of evolution through the process of learning.

The most common principle behind learning is the Law of Primary Reinforcement: the tendency of contiguous activities associated with the termination of needs to be reinvoked on similar occasions. This principle can be applied not only to simple phenomena such as the Pavlovian conditioned response but also to the explanation of more complex selective learning.

By constantly referring to empirical experimentation the author derives sixteen basic postulates having to do with the factors of reinforcement, generalization, motivation, inhibition, variation, and the like. Most of these factors are symbolic constructs and the condition of their acceptance is always the pragmatic test of agreement with and explanation of real events.

By gathering together his theories in one volume the author has performed a great service to theoretical psychology. His emphasis on mathematical formulation and rigid definition is a major step in the attempt to place the science of behavior on a par with the 'exact' sciences. In no sense is a claim made that the surface of the problem has been any more than scratched. It may be noted, indeed, that the relevancies of the system to everyday behavior are sometimes remote. Nevertheless, many psychologists will agree that the theoretical psychology of the future lies in the direction of the type of approach made in this book.

LEONARD S. KOGAN

University of Rochester

CORRECTION

In May, 1942, issue of this *Journal* certain errors were made in printing the article, "A Ratio for Estimating the Reliability of Test Scores," by John M. Butler. The following corrections should be made:

In lines three and eight, page 392, the expression, $s\sqrt{\frac{1 - r_{tt}}{s}}$ should be $\frac{s\sqrt{1 - r_{tt}}}{s}$. Lines nine and ten, page 393, should read "Since r_{tt} and $1 - r_{tt}$ are variances, it is obvious that $F = \frac{r_{tt}}{1 - r_{tt}}$ when $r_{tt} > .50$, and $F = \frac{1 - r_{tt}}{r_{tt}}$ when $r_{tt} < .50$." instead of "Since r_{tt} and $1 - r_{tt}$ are variances, it is obvious that $F = \frac{r_{tt}}{1 - r_{tt}}$ when r_{tt} .50 & $F = \frac{1 - r_{tt}}{r_{tt}}$ when r_{tt} .50".

AMERICAN PSYCHOLOGICAL PERIODICALS

American Journal of Psychology—Ithaca, N.Y.; Cornell University. Subscription $6.50. 624 pages annually. Edited by K. M. Dallenbach, Madison Bentley, and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887

Journal of Genetic Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 1000 pages annually. Edited by Carl Murchison. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Psychological Review—Northwestern University, Evanston, Ill.; American Psychological Association, Inc. Subscription $5.50. 540 pages annually. Edited by Herbert S. Langfeld. Bi-monthly General psychology. Founded 1894.

Psychological Monographs—Northwestern University, Evanston, Ill.; American Psychological Association, Inc. Subscription $6.00 per volume. 500 pages. Edited by John F. Dashiell. Without fixed dates, each issue one or more researches. Founded 1895.

Psychological Bulletin—Northwestern University, Evanston, Ill.; American Psychological Association, Inc. Subscription $7.00. 665 pages annually. Edited by John A. McGeoch. Monthly (10 numbers). Psychological literature. Founded 1904.

Archives of Psychology—New York, N.Y.; Columbia University. Subscription $6.00 per volume. 500 pages. Edited by R. S. Woodworth. Without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Northwestern University, Evanston, Ill.; American Psychological Association, Inc. Subscription $5.00. 560 pages annually. Edited by Gordon W. Allport. Quarterly. Founded 1906.

Journal of Educational Psychology—Baltimore, Md.; Warwick & York. Subscription $6.00. 576 pages annually. J. W. Dunlap, Chairman of the Editorial Board. Monthly except June to August. Founded 1910.

Psychoanalytic Review—New York, N.Y.; 64 West 56th St. Subscription $6.00. 500 pages annually Edited by Smith Ely Jelliffe. Quarterly. Founded 1913.

Journal of Experimental Psychology—Northwestern University, Evanston, Ill.; American Psychological Association, Inc. Subscription $14.00 per annum (2 volumes). 1250 pages annually. Edited by Samuel W. Fernberger. Monthly. Founded 1916.

Journal of Applied Psychology—Indianapolis, Ind.; C. E. Pauley & Co. Subscription $6.00. 600 pages annually. Edited by James P. Porter. Bi-monthly. Founded 1917.

Journal of Comparative Psychology—Baltimore, Md.; Williams & Wilkins Co. Subscription $10.00 per annum (2 volumes). 900 pages annually. Edited by Roy M. Dorcus, Knight Dunlap, and Robert M. Yerkes. Bi-monthly. Founded 1921.

Comparative Psychology Monographs—Baltimore, Md.; The Johns Hopkins Press. Subscription $5.00 per volume. 400 pages. Edited by Roy M. Dorcus. Without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Provincetown, Mass.; The Journal Press. Subscription $7.00. 500 pages annually. Edited by Carl Murchison. Bi-monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

Psychological Abstracts—Northwestern University, Evanston, Ill.; American Psychological Association, Inc. Subscription $7.00. 700 pages annually. Edited by Walter S. Hunter and H. L. Ansbacher. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 1000 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Provincetown, Mass.; The Journal Press. Subscription $7.00. 500 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Psychoanalytic Quarterly—Albany, N.Y.; 372–374 Broadway. Subscription $6.00. 560 pages annually Edited by Bertram D. Lewin and others. Quarterly. Founded 1932.

Character and Personality—Durham, N.C.; Duke University Press. Subscription $2.00. 360 pages annually. Edited by Karl Zener. Quarterly. Founded 1932.

Journal of Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 800–1200 pages annually. Edited by Carl Murchison. Quarterly. Founded 1936.

Psychometrika—University of Chicago, Ill.; Psychometric Society. Subscription $10.00. 320 pages annually. Edited by L. L. Thurstone and others. Quarterly. Quantitative methods in psychology. Founded 1936.

Psychological Record—Bloomington, Ind.; Principia Press. Subscription $4.00. 500 pages annually. Edited by J. R. Kantor and C. M. Louttit. Without fixed dates, each number a single research. General psychology. Founded 1937.

Journal of Consulting Psychology—Lancaster, Penn.; Science Printing Co. Subscription $3.00. 240 pages annually. Edited by J. P. Symonds. Bi-monthly. Founded 1937.

Forthcoming Macmillan Books

PRACTICAL PSYCHOLOGY

By F. K. Berrien

Assistant Professor of Psychology, Colgate University

This new book surveys the principal problems and outstanding findings in the field of practical psychology. It includes several chapters on subjects not ordinarily included in books of this type—detecting deception covering experimental evidence, legal status, and practical applications of the techniques; effective public speaking and writing; and morale and rewards for work. The author is an experienced practicing psychologist as well as a teacher of psychology.

Ready in Feb. Ill. $4.00 (*probable*)

Introduction to EXCEPTIONAL CHILDREN

By Harry J. Baker

Director of the Psychological Clinic, Detroit Public Schools

The thirty chapters which comprise this new book examine all aspects of exceptional children—vision, hearing, speech, crippled conditions, neurological and psychogenic disorders, behavior problems, etc. Children with exceptional minds as well as physically handicapped children are discussed. Each abnormality is described, interpreted and related to the total adjustment of the child.

Ready in Jan. Ill. $3.50 (*probable*)

The Macmillan Company · 60 Fifth Avenue · New York 11

THE ACTIVITY MOVEMENT

By CLYDE HISSONG

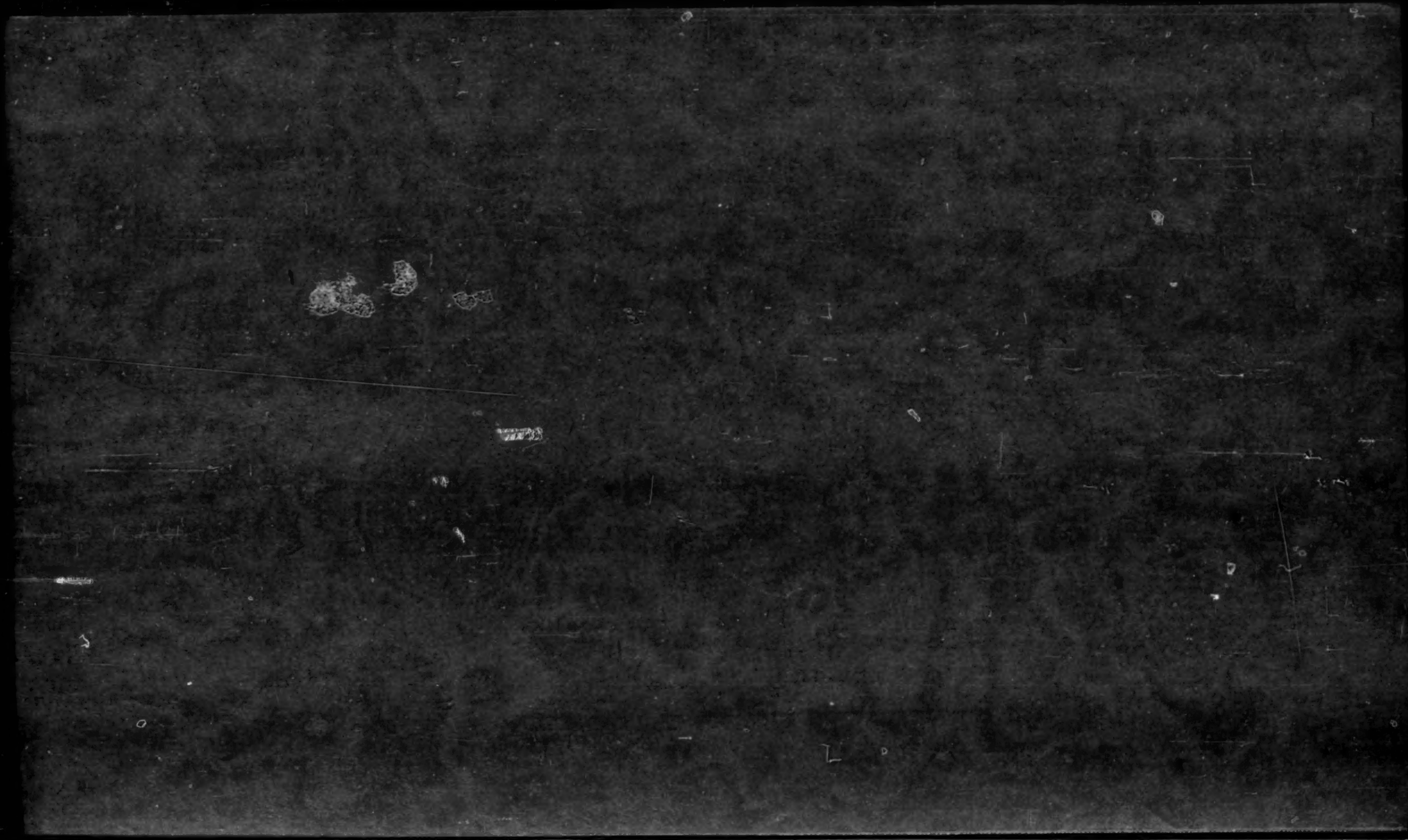
In an attempt to overcome the weaknesses of the traditional school organization many progressive schools have developed new programs. These programs are so similar in character that collectively the changes have been referred to as the activity movement. This movement has claimed the center of the educational stage for a length of time sufficient to have engendered widespread interest in its outcomes and in its basic philosophy.

In Doctor Hissong's study an attempt has been made to discover the principles underlying the present activity movement, to determine the influence of traditional concepts in shaping the trends of the movement, and to see if in the light of the present knowledge of the child and his relation to his environment the movement rests upon a justifiable basis.

$2.00 plus 10¢ postage

WARWICK AND YORK

Publishers BALTIMORE

THE MAPLE PRESS COMPANY, YORK, PA.

INCHES
1
2
3
4
5
6
C2399
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
METRIC